Final value problems for parabolic differential equations and their well-posedness

Christensen, Ann-Eva; Johnsen, Jon

Published in:
Axioms

DOI (link to publication from Publisher):
10.3390/axioms7020031

Publication date:
2018

Document Version
Publisher's PDF, also known as Version of record

axioms

MDPI

Article

Final Value Problems for Parabolic Differential Equations and Their Well-Posedness

Ann-Eva Christensen [1,†] and Jon Johnsen [2,*,†]

[1] Unit of Epidemiology and Biostatistics, Aalborg University Hospital, Hobrovej 18-22, DK-9000 Aalborg, Denmark; anneva.ch@gmail.com

[2] Department of Mathematics, Aalborg University, Skjernvej 4A, DK-9220 Aalborg Øst, Denmark

* Correspondence: jjohnsen@math.aau.dk; Tel.: +45-9940-8847

† These authors contributed equally to this work.

Received: 29 March 2018; Accepted: 28 April 2018; Published: 9 May 2018

check for updates

Abstract: This article concerns the basic understanding of parabolic final value problems, and a large class of such problems is proved to be well posed. The clarification is obtained via explicit Hilbert spaces that characterise the possible data, giving existence, uniqueness and stability of the corresponding solutions. The data space is given as the graph normed domain of an unbounded operator occurring naturally in the theory. It induces a new compatibility condition, which relies on the fact, shown here, that analytic semigroups always are invertible in the class of closed operators. The general set-up is evolution equations for Lax–Milgram operators in spaces of vector distributions. As a main example, the final value problem of the heat equation on a smooth open set is treated, and non-zero Dirichlet data are shown to require a non-trivial extension of the compatibility condition by addition of an improper Bochner integral.

Keywords: parabolic boundary problem; final value; compatibility condition; well posed; non-selfadjoint; hyponormal

MSC: 35A01; 47D06

---

1. Introduction

In this article, we establish well-posedness of final value problems for a large class of parabolic differential equations. Seemingly, this clarifies a longstanding gap in the comprehension of such problems.

Taking the heat equation as a first example, we address the problem of characterising the functions $u(t,x)$ that, in a C^∞-smooth bounded open set $\Omega \subset \mathbb{R}^n$ with boundary $\partial\Omega$, fulfil the equations, where $\Delta = \partial_{x_1}^2 + \cdots + \partial_{x_n}^2$ denotes the Laplacian,

$$\begin{cases} \partial_t u(t,x) - \Delta u(t,x) = f(t,x) & \text{for } t \in \,]0,T[\,,\ x \in \Omega, \\ u(t,x) = g(t,x) & \text{for } t \in \,]0,T[\,,\ x \in \partial\Omega, \\ u(T,x) = u_T(x) & \text{for } x \in \Omega. \end{cases} \tag{1}$$

Motivation for doing so could be given by imagining a nuclear power plant, which is hit by a power failure at time $t = 0$. Once power is regained at time $t = T$, and a measurement of the reactor temperature $u_T(x)$ is obtained, it is of course desirable to calculate backwards in time to provide an answer to the question: were temperatures $u(t,x)$ around some earlier time $t_0 < T$ high enough to cause a meltdown of the fuel rods ?

We provide here a theoretical analysis of such problems and prove that they are well-posed, that is, they have *existence, uniqueness* and *stability* of solutions $u \in X$ for given data $(f, g, u_T) \in Y$, in certain normed spaces X, Y to be specified below. The results were announced without proofs in the short note [1].

Although well-posedness is of decisive importance for the interpretation and accuracy of numerical schemes, which one would use in practice, such a theory has seemingly not been worked out before. Explained roughly, our method is to provide a useful structure on the reachable set for a general class of parabolic differential equations.

1.1. Background

Let us first describe the case $f = 0$, $g = 0$. Then the mere heat equation $(\partial_t - \Delta)u = 0$ is clearly solved for all $t \in \mathbb{R}$ by the function $u(t,x) = e^{(T-t)\lambda}v(x)$, if $v(x)$ is an eigenfunction of the Dirichlet realization $-\Delta_D$ of the Laplace operator with eigenvalue λ.

In view of this, the homogeneous final value problem (1) would obviously have the above u as a *basic* solution if, coincidentally, the final data $u_T(x)$ were given as the eigenfunction $v(x)$. The theory below includes the set $\mathcal{B}$ of such basic solutions u together with its linear hull $\mathcal{E} = \operatorname{span}\mathcal{B}$ and a certain completion $\overline{\mathcal{E}}$.

It is easy to describe $\mathcal{E}$ in terms of the eigenvalues $0 < \lambda_1 \le \lambda_2 \le \dots$ and the associated $L_2(\Omega)$-orthonormal basis $e_1, e_2, \dots$ of eigenfunctions of $-\Delta_D$: corresponding to final data u_T in $\operatorname{span}(e_j)$, which are the u_T having *finite* expansions $u_T(x) = \sum_j (u_T \mid e_j) e_j(x)$ in $L_2(\Omega)$, the space $\mathcal{E}$ consists of solutions $u(t,x)$ being *finite* sums

$$u(t,x) = \textstyle\sum_j e^{(T-t)\lambda_j}(u_T \mid e_j)e_j(x). \tag{2}$$

Moreover, at time $t = 0$ there is, because of the finiteness, a vector $u(0,x)$ in $L_2(\Omega)$ that trivially fulfills

$$\|u(0,\cdot)\|^2 = \textstyle\sum_j e^{2T\lambda_j}|(u_T \mid e_j)|^2 < \infty. \tag{3}$$

However, when summation is extended to all $j \in \mathbb{N}$, condition (3) becomes very strong, as it is only satisfied for special u_T: Weyl's law for the counting function, cf. ([2], Chapter 6.4), entails the well-known $\lambda_j = \mathcal{O}(j^{\frac{2}{n}})$, so a single term in (3) yields $|(u_T \mid e_j)| \le c\exp(-Tj^{\frac{2}{n}})$; i.e., the L_2-coordinates of such u_T decay rapidly for $j \to \infty$.

Condition (3) has been known at least since the 1950s; the work of John [3] and Miranker [4] are the earliest we know. While many authors have avoided an analysis of it, Payne found it scientifically intolerable because u_T is likely to be imprecisely measured; cf. his treatise [5] on the variety of methods applied to (1) until the mid 1970s.

More recently, e.g., Isakov [6] emphasized the classical observation, found already in [4], that (2) implies a phenomenon of *instability*. Indeed, the sequence of final data $u_{T,k} = e_k$ has constant length 1, yet via (2) it gives the initial states $u_k(0,x) = e^{T\lambda_k}e_k(x)$ having L_2-norms $e^{T\lambda_k}$, which clearly blow up rapidly for $k \to \infty$.

This L_2-instability cannot be explained away, of course, but it does *not* rule out that (1) is well-posed. It rather indicates that the L_2-norm is an insensitive choice for (1).

In fact, here there is an analogy with the classical stationary Dirichlet problem

$$-\Delta u = f \quad \text{in } \Omega, \qquad u = g \quad \text{on } \partial\Omega. \tag{4}$$

This is *unsolvable* for $u \in C^2(\overline{\Omega})$ given certain $f \in C^0(\overline{\Omega})$, $g \in C^0(\partial\Omega)$: Günther proved prior to 1934, cf. ([7], p. 85), that when $f(x) = \chi(x)(3x_3^2|x|^{-2} - 1)/\log|x|$ for some radial cut-off function $\chi \in C_0^\infty(\Omega)$ equal to 1 around the origin, Ω being the unit ball of $\mathbb{R}^3$, then $f \in C^0(\overline{\Omega})$ while the convolution $w = \frac{1}{4\pi|x|} * f$ is in $C^1(\overline{\Omega})$ but *not* in C^2 at $x = 0$; so $w \in C^1(\overline{\Omega}) \setminus C^2(\overline{\Omega})$. Yet w is the unique $C^1(\overline{\Omega})$-solution of (4) in the distribution space $\mathcal{D}'(\Omega)$ in the case g is given as $g = w_{|\partial\Omega}$.

Thus the C^k-scales constitute an insensitive choice for (4). Nonetheless, replacing $C^2(\overline{\Omega})$ by its completion $H^1(\Omega)$ in the Sobolev norm $(\sum_{|\alpha|\leq 1}\int_\Omega |D^\alpha u|^2\,dx)^{1/2}$, it is classical that (4) is well-posed with u in $H^1(\Omega)$.

To obtain similarly well-adapted spaces for (1) with $f=0$, $g=0$, one could base the analysis on (3). Indeed, along with the above space $\mathcal{E}$ of basic solutions, a norm $|||u_T|||$ on the space of final data $u_T \in \operatorname{span}(e_j)$ can be *defined* by (3), leading to the norm $|||u_T||| = (\sum_{j=1}^\infty e^{2T\lambda_j}|(u_T\,|\,e_j)|^2)^{1/2}$ on the u_T that correspond to solutions u in the completion $\overline{\mathcal{E}}$. This would give well-posedness of (1) with $u \in \overline{\mathcal{E}}$; cf. Remark 16.

But the present paper goes much beyond this. For one thing, we have freed the discussion from $-\Delta_D$ and its specific eigenvalue distribution by using sesqui-linear forms, cf. Lax–Milgram's lemma, which allowed us to extend the proofs to a general class of elliptic operators A.

Secondly we analyse the *fully* inhomogeneous problem (1) for general f, g in Section 5. In this situation well-posedness is not just a matter of choosing the norm on the data (f,g,u_T) suitably (as one might think from the above $|||u_T|||$). In fact, prior to this choice, one has to *restrict* the (f,g,u_T) to a subspace characterised by certain *compatibility conditions*. While such conditions are well known in the theory of parabolic boundary problems, they are shown here to have a new and special form for final value problems.

Indeed, the compatibility conditions stem from the unbounded operator $u_T \mapsto u(0)$, which maps the final data to the corresponding initial state in the presence of the source term f. The fact that this operator is well defined, and that its domain endowed with the graph norm yields the data space, is the leitmotif of this article.

1.2. The Abstract Final Value Problem

Let us outline our analysis given for a Lax–Milgram operator A defined in H from a V-elliptic sesquilinear form $a(\cdot,\cdot)$ in a Gelfand triple, i.e., in a set-up of three Hilbert spaces $V \hookrightarrow H \hookrightarrow V^*$ having norms denoted $\|\cdot\|$, $|\cdot|$ and $\|\cdot\|_*$, and where V is the form domain of a.

In this framework, we consider the following general final value problem: given data

$$f \in L_2(0,T;V^*), \qquad u_T \in H, \tag{5}$$

determine the V-valued vector distributions $u(t)$ on $]0,T[$, that is the $u \in \mathcal{D}'(0,T;V)$, fulfilling

$$\begin{cases} \partial_t u + Au = f & \text{in } \mathcal{D}'(0,T;V^*), \\ \quad u(T) = u_T & \text{in } H. \end{cases} \tag{6}$$

Classically a wealth of parabolic Cauchy problems with homogeneous boundary conditions have been efficiently treated with the triples (H,V,a) and the $\mathcal{D}'(0,T;V^*)$ set-up in (6). For this the reader may consult the work of Lions and Magenes [8], Tanabe [9], Temam [10], Amann [11]. Also recently, e.g., Almog, Grebenkov, Helffer, Henry studied variants of the complex Airy operator via such triples [12–14], and our results should at least extend to final value problems for those of their realisations that have non-empty spectrum.

To compare (6) with the analogous Cauchy problem, we recall that whenever $u' + Au = f$ is solved under the initial condition $u(0) = u_0 \in H$, for some $f \in L_2(0,T;V^*)$, there is a unique solution u in the Banach space

$$\begin{aligned} X =& L_2(0,T;V) \bigcap C([0,T];H) \bigcap H^1(0,T;V^*), \\ \|u\|_X =& \Big(\int_0^T \|u(t)\|^2\,dt + \sup_{0\leq t\leq T}|u(t)|^2 + \int_0^T (\|u(t)\|_*^2 + \|u'(t)\|_*^2)\,dt\Big)^{1/2}. \end{aligned} \tag{7}$$

For (6) it would thus be natural to envisage solutions u in the same space X. This turns out to be true, but only under substantial further conditions on the data (f,u_T).

To formulate these, we exploit that $-A$ generates an *analytic* semigroup e^{-tA} in $\mathbb{B}(H)$. This is crucial for the entire article, because analytic semigroups always are invertible in the class of closed operators, as we show in Proposition 1. We denote its inverse by e^{tA}, consistent with the case $-A$ generates a group,

$$(e^{-tA})^{-1} = e^{tA}. \tag{8}$$

Its domain is the Hilbert space $D(e^{tA}) = R(e^{-tA})$ that is normed by $\|u\| = (|u|^2 + |e^{tA}u|^2)^{1/2}$. In Proposition 10 we show that a non-empty spectrum, $\sigma(A) \neq \emptyset$, yields strict inclusions

$$D(e^{t'A}) \subsetneq D(e^{tA}) \subsetneq H \qquad \text{for } 0 < t < t'. \tag{9}$$

For $t = T$ these domains play a crucial role in the well-posedness result below, cf. (11), where also the full yield y_f of the source term f on the system appears, namely

$$y_f = \int_0^T e^{-(T-s)A} f(s)\, ds. \tag{10}$$

The map $f \mapsto y_f$ takes values in H, and it is a continuous surjection $y_f\colon L_2(0,T;V^*) \to H$.

Theorem 1. *For the final value problem (6) to have a solution u in the space X in (7), it is necessary and sufficient that the data (f, u_T) belong to the subspace Y of $L_2(0,T;V^*) \oplus H$ defined by the condition*

$$u_T - \int_0^T e^{-(T-t)A} f(t)\, dt \;\in\; D(e^{TA}). \tag{11}$$

Moreover, in X the solution u is unique and depends continuously on the data (f, u_T) in Y, that is, we have $\|u\|_X \leq c\|(f, u_T)\|_Y$, when Y is given the graph norm

$$\|(f, u_T)\|_Y = \left(|u_T|^2 + \int_0^T \|f(t)\|_*^2\, dt + \left| e^{TA}\left(u_T - \int_0^T e^{-(T-t)A} f(t)\, dt\right) \right|^2 \right)^{1/2}. \tag{12}$$

(The full statements are found in Theorems 7 and 8 below.)

Condition (11) is a fundamental novelty for the above class of final value problems, but more generally it also gives an important clarification for parabolic differential equations.

As for its nature, we note that the data (f, u_T) fulfilling (11) form a Hilbert(-able) space Y embedded into $L_2(0,T;V^*) \oplus H$, in view of its norm in (12).

Using the above y_f, (12) is seen to be the graph norm of $(f, u_T) \mapsto e^{TA}(u_T - y_f)$, which in terms of $\Phi(f, u_T) = u_T - y_f$ is the unbounded operator $e^{TA}\Phi$ from $L_2(0,T;V^*) \oplus H$ to H. As (11) means that the operator $e^{TA}\Phi$ must be defined at (f, u_T), the space Y is its domain. Thus $e^{TA}\Phi$ is a key ingredient in the rigorous treatment of (6).

The role of $e^{TA}\Phi$ is easy to elucidate in control theoretic terms: its value $e^{TA}\Phi(f, u_T)$ simply equals the particular initial state $u(0)$ which is steered by f to the final state $u(T) = u_T$ at time T; cf. (13) below.

Because of $e^{-(T-t)A}$ and the integral over $[0, T]$, (11) involves *non-local* operators in both space and time as an inconvenient aspect—exacerbated by use of the abstract domain $D(e^{TA})$, which for longer lengths T of the time interval gives increasingly stricter conditions; cf. (9).

Anyhow, we propose to regard (11) as a *compatibility* condition on the data (f, u_T), and thus we generalise the notion of compatibility.

For comparison we recall that Grubb and Solonnikov [15] made a systematic investigation of a large class of *initial*-boundary problems of parabolic (pseudo-)differential equations and worked out compatibility conditions, which are necessary and sufficient for well-posedness in full scales of anisotropic L_2-Sobolev spaces. Their conditions are explicit and local at the curved corner $\partial\Omega \times \{0\}$, except for half-integer values of the smoothness s that were shown to require so-called coincidence, which is expressed in integrals over the product of the two boundaries $\{0\} \times \Omega$ and $]0, T[\times \partial\Omega$; hence it also is a non-local condition.

However, while the conditions of Grubb and Solonnikov [15] are decisive for the solution's regularity, condition (11) is crucial for the *existence* question; cf. the theorem.

Previously, uniqueness was shown by Amann ([11], Section V.2.5.2) in a t-dependent set-up, but injectivity of $u(0) \mapsto u(T)$ was proved much earlier for problems with t-dependent sesquilinear forms by Lions and Malgrange [16].

Showalter [17] attempted to characterise the possible u_T in terms of Yosida approximations for $f = 0$ and A having half-angle $\frac{\pi}{4}$. As an ingredient, invertibility of analytic semigroups was claimed in [17] for such A, but the proof was flawed as A can have semi-angle $\pi/4$ even if A^2 is not accretive; cf. our example in Remark 9.

Theorem 1 is proved largely by comparing with the corresponding problem $u' + Au = f$, $u(0) = u_0$. It is well known in functional analysis, cf. (7), that this is well-posed for $f \in L_2(0, T; V^*)$, $u_0 \in H$, with solutions $u \in X$. However, as shown below by adaptation of a classical argument, u is also in this set-up necessarily given by Duhamel's principle, or the variation of constants formula, for the analytic semigroup e^{-tA} in V^*,

$$u(t) = e^{-tA}u(0) + \int_0^t e^{-(t-s)A} f(s)\, ds. \tag{13}$$

For $t = T$ this yields a *bijective correspondence* $u(0) \leftrightarrow u(T)$ between the initial and terminal states (in particular backwards uniqueness of the solutions in the large class X)—but this relies crucially on the previously mentioned invertibility of e^{-tA}; cf. (8).

As a consequence of (13) one finds the necessity of (11), as the difference $\Phi(f, u_T) = u_T - y_f$ in (11) must equal the vector $e^{-TA}u(0)$, which obviously belongs to $D(e^{TA})$.

Moreover, (13) yields that $u(T)$ in a natural way consists of two parts, that differ radically even when A has nice properties:

First, $e^{-tA}u(0)$ solves the semi-homogeneous problem with $f = 0$, and for $u(0) \neq 0$ there is the precise property in non-selfadjoint dynamics that the "height" function $h(t)$ is *strictly convex*,

$$h(t) = |e^{-tA}u(0)|. \tag{14}$$

This is shown in Proposition 4 when A belongs to the broad class of hyponormal operators, studied by Janas [18], or in case A^2 is accretive; then $h(t)$ is also strictly decreasing with $h'(0) \leq -m(A)$, where $m(A)$ is the lower bound of A.

The stiffness inherent in strict convexity is supplemented by the fact that $u(T) = e^{-TA}u(0)$ is confined to a dense, but very small space, as by a well-known property of analytic semigroups,

$$u(T) \in \textstyle\bigcap_{n\in\mathbb{N}} D(A^n). \tag{15}$$

Secondly, for $u_0 = 0$ the integral in (13) solves the initial value problem, and it has a rather different nature since its final value y_f in (10) is surjective $y_f \colon L_2(0, T; V^*) \to H$, hence can be *anywhere* in H, regardless of the Lax–Milgram operator A in our set-up. This we show in Proposition 6 using a kind of control-theoretic argument in case A is self-adjoint with compact inverse; and for general A by means of the Closed Range Theorem, cf. Proposition 5.

For the reachable set of the equation $u' + Au = f$, or rather the possible final data u_T, they will be a sum of an arbitrary vector y_f in H and a term $e^{-TA}u(0)$ of great stiffness (cf. (15)). Thus u_T can be prescribed in the affine space $y_f + D(e^{TA})$. As any $y_f \neq 0$ will push the dense set $D(e^{TA}) \subset H$ in some arbitrary direction, $u(T)$ can be expected *anywhere* in H (unless $y_f \in D(e^{TA})$ is known a priori). Consequently neither $u(T) \in D(e^{TA})$ nor (15) can be expected to hold for $y_f \neq 0$, not even if its norm $|y_f|$ is much smaller than $|e^{-TA}u(0)|$.

As for final state measurements in real life applications, we would like to prevent a misunderstanding by noting that it is only under the peculiar circumstance that $y_f = 0$ is known *a priori* to be an *exact* identity that (15) would be a valid expectation on $u(T)$.

Indeed, even if f is so small that it is (quantitatively) insignificant for the time development of the system governed by $u' + Au = f$, so that $f = 0$ is a valid dynamical approximation, the (qualitative) *mathematical* expectation that $u(T)$ should fulfill (15) cannot be justified from such an approximation; cf. the above.

In view of this fundamental difference between the problems that are truly and merely approximately homogeneous, it seems that proper understanding of final value problems is facilitated by treating inhomogeneous problems from the very beginning.

1.3. The Inhomogeneous Heat Problem

For (1) with general data (f, g, u_T) the above is applied with $A = -\Delta_D$, that is the Dirichlet realisation of the Laplacian. The results are analogous, but less simple to state and more demanding to obtain.

First of all, even though it is a linear problem, the compatibility condition (11) *destroys* the old trick of reducing to boundary data $g = 0$, for when $w \in H^1$ fulfils $w = g \neq 0$ on the curved boundary $S =]0, T[\times \partial\Omega$, then w *lacks* the regularity needed to test (11) on the data $(\tilde{f}, 0, \tilde{u}_T)$ of the reduced problem; cf. (127) ff.

Secondly, it is, therefore, non-trivial to clarify that every $g \neq 0$ *does* give rise to an extra term z_g, in the sense that (11) is replaced by the compatibility condition

$$u_T - y_f + z_g \in D(e^{-T\Delta_D}). \tag{16}$$

Thirdly, due to the low reqularity, it requires technical diligence to show that z_g, despite the singularity of $\Delta e^{(T-s)\Delta_D}$ at $s = T$, has the structure of a single convergent improper Bochner integral, namely

$$z_g = \oint_0^T \Delta e^{(T-s)\Delta_D} K_0 g(s)\, ds. \tag{17}$$

The reader is referred to Section 5 for the choice of the Poisson operator K_0 and for an account of the results on the fully inhomogeneous problem in (1), especially Theorem 10 and Corollary 3, which we sum up here:

Theorem 2. *For given data $f \in L_2(0, T; H^{-1}(\Omega))$, $g \in H^{1/2}(S)$, $u_T \in L_2(\Omega)$ the final value problem (1) is solved by a function u in $X_1 = L_2(0, T; H^1(\Omega)) \bigcap C([0, T]; L_2(\Omega)) \bigcap H^1(0, T; H^{-1}(\Omega))$, if and only if the data in terms of (10) and (17) satisfy the compatibility condition (16). In the affirmative case, u is uniquely determined in X_1 and has the representation, with all terms in X_1,*

$$u(t) = e^{t\Delta_D} e^{-T\Delta_D}(u_T - y_f + z_g) + \int_0^t e^{(t-s)\Delta} f(s)\, ds - \oint_0^t \Delta e^{(t-s)\Delta_D} K_0 g(s)\, ds, \tag{18}$$

The unique solution u in X_1 depends continuously on the data (f, g, u_T) in the Hilbert space Y_1, when these are given the norms in (130) and (158) below, respectively.

1.4. Contents

Our presentation is aimed at describing methods and consequences in a concise way, readable for a broad audience within evolution problems. Therefore we have preferred a simple set-up, leaving many examples and extensions to future work, cf. Section 6.

Notation is given in Section 2 together with the set-up for Lax–Milgram operators and semigroup theory. Some facts on forward evolution problems are recalled in Section 3, followed by our analysis of abstract final value problems in Section 4. The heat equation and its final and boundary value problems are treated in Section 5. Section 6 concludes with remarks on the method's applicability and notes on the literature of the problem.

2. Preliminaries

In the sequel specific constants will appear as C_j, $j \in \mathbb{N}$, whereas constants denoted by c may vary from place to place. $\mathbf{1}_S$ denotes the characteristic function of the set S.

Throughout V and H denote two separable Hilbert spaces, such that V is algebraically, topologically and densely contained in H. Then there is a similar inclusion into the anti-dual V^*, i.e., the space of conjugated linear functionals on V,

$$V \subseteq H \equiv H^* \subseteq V^*. \tag{19}$$

(V, H, V^*) is also known as a Gelfand triple. Denoting the norms by $\|\cdot\|$, $|\cdot|$ and $\|\cdot\|_*$, respectively, there are constants such that for all $v \in V$,

$$\|v\|_* \leq C_1|v| \leq C_2\|v\|. \tag{20}$$

The inner product on H is denoted by $(\cdot\,|\,\cdot)$; and the sesquilinear scalar product on $V^* \times V$ by $\langle\cdot,\cdot\rangle_{V^*,V}$ or $\langle\cdot,\cdot\rangle$, it fulfils $|\langle u, v\rangle| \leq \|u\|_*\|v\|$. The second inclusion in (19) means that for $u \in H$,

$$\langle u, v\rangle = (u\,|\,v) \quad \text{ for all } v \in V. \tag{21}$$

For a linear transformation A in H, the domain is written $D(A)$, while $R(A)$ denotes its range and $Z(A)$ its null-space. $\rho(A)$, $\sigma(A)$ and $\nu(A) = \{\, (Au\,|\,u) \mid u \in D(A),\ |u| = 1 \,\}$ denote the resolvent set, spectrum and numerical range, while $m(A) = \inf \operatorname{Re} \nu(A)$ is the lower bound of A. Throughout $\mathbb{B}(H)$ stands for the Banach space of bounded linear operators on H.

For a given Banach space B and $T > 0$, we denote by $L_1(0, T; B)$ the space of equivalence classes of functions $f\colon [0, T] \to B$ that are strongly measurable with $\int_0^T \|f(t)\|\,dt$ finite. For such f the Bochner integral is denoted by $\int_0^T f(t)\,dt$, cf. [19]; it fulfils $\langle \int_0^T f(t)\,dt,\ \lambda \rangle = \int_0^T \langle f(t),\ \lambda \rangle\,dt$ for every functional λ in the dual space B'. Likewise $L_2(0, T, B)$ consists of the strongly measurable f with finite norm $(\int_0^T \|f(t)\|^2\,dt)^{1/2}$.

On an open set $\Omega \subset \mathbb{R}^n$, $n \geq 1$, the space $C_0^\infty(\Omega)$ consists of the infinitely differentiable functions having compact support in Ω; it is given the usual $\mathcal{LF}$-topology, cf. [20,21]. The dual space of continuous linear functionals $\mathcal{D}'(\Omega)$ is the distribution space on Ω. We use the standard distribution theory as exposed by Grubb [20] and Hörmander [22].

More generally, the space of B-valued vector distributions is denoted by $\mathcal{D}'(\Omega; B)$; it consists of the continuous linear maps $\Lambda\colon C_0^\infty(\Omega) \to B$, cf. [21], the value of which at $\varphi \in C_0^\infty(\Omega)$ is indicated by $\langle \Lambda,\ \varphi \rangle$. If Ω is the interval $]0, T[$ we also write $\mathcal{D}'(\Omega; B) = \mathcal{D}'(0, T; B)$.

The Sobolev space $H^1(0, T; B)$ consists of the $u \in \mathcal{D}'(0, T; B)$ for which both u, u' belong to $L_2(0, T; B)$; it is normed by $\int_0^T (\|u\|^2 + \|u'\|^2)\,dt)^{1/2}$. More generally $W^{1,1}(0, T; B)$ is defined by replacing L_2 by L_1.

2.1. Lax–Milgram Operators

Our main tool will be the Lax–Milgram operator associated to an elliptic sesquilinear form, cf. the set-up in ([20], Section 12.4). For the reader's sake we review this, also to establish a few additional points from the proofs in [20].

We let $a(\cdot, \cdot)$ be a bounded, V-elliptic sesquilinear form on V, i.e., certain $C_3, C_4 > 0$ fulfil for all $u, v \in V$

$$|a(u, v)| \leq C_3\|u\|\|v\|, \qquad \operatorname{Re} a(v, v) \geq C_4\|v\|^2. \tag{22}$$

Obviously, the adjoint sesquilinear form $a^*(u,v) = \overline{a(v,u)}$ inherits these properties (with the same C_3, C_4), and so does the "real part", $a_{\mathrm{Re}}(u,v) = \frac{1}{2}(a(u,v) + a^*(u,v))$. Since $a_{\mathrm{Re}}(u,u) \geq 0$, the form a_{Re} is an inner product on V, inducing the equivalent norm

$$|||u||| = a_{\mathrm{Re}}(u,u)^{1/2}, \quad \text{for } u \in V. \tag{23}$$

We recall that $s(u,v) = (Su \,|\, v)_V$ gives a bijective correspondence between bounded sesquilinear forms $s(\cdot,\cdot)$ on V and bounded operators $S \in \mathbb{B}(V)$, which is isometric since $\|S\|$ equals the operator norm of the sesquilinear form $|s| = \sup\{|s(u,v)| \mid \|u\| = 1 = \|v\|\}$. So the given form a induces an $\mathcal{A}_0 \in \mathbb{B}(V)$ given by

$$a(u,v) = (\mathcal{A}_0 u \,|\, v)_V \quad \forall u,v \in V; \tag{24}$$

and the adjoint form a^* similarly induces an operator $\mathcal{A}_0^* \in \mathbb{B}(V)$, which is seen at once to be the adjoint of $\mathcal{A}_0$ in the sense that $(\mathcal{A}_0^* v \,|\, u)_V = (v \,|\, \mathcal{A}_0 u)_V$.

The V-ellipticity in (22) shows that $\mathcal{A}_0$, $\mathcal{A}_0^*$ are both injective with positive lower bounds $m(\mathcal{A}_0)$, $m(\mathcal{A}_0^*) \geq C_4$, so $\mathcal{A}_0$, $\mathcal{A}_0^*$ are in fact bijections on V (cf. ([20], Theorem 12.9)).

By Riesz's representation theorem, there exists a bijective isometry $J \in \mathbb{B}(V, V^*)$ such that for every $v^* = J\tilde{v}$ one has $\langle J\tilde{v}, v\rangle = (\tilde{v} \,|\, v)_V$ for all $v \in V$. Therefore $\mathcal{A} := J \circ \mathcal{A}_0$ is an operator in $\mathbb{B}(V, V^*)$, for which (24) gives

$$\langle \mathcal{A}u, v\rangle = a(u,v), \quad \forall u,v \in V. \tag{25}$$

Similarly $\mathcal{A}' := J \circ \mathcal{A}_0^*$ fulfils $\langle \mathcal{A}'u, v\rangle = (\mathcal{A}_0^* u \,|\, v)_V = a^*(u,v)$ for all $u,v \in V$.

Clearly $\mathcal{A}$ and $\mathcal{A}'$ are bijections, as composites of such. Hence they give rise to a Hilbert space structure on V^* with the inner product

$$(w_1 \,|\, w_2)_{V^*} = a_{\mathrm{Re}}(\mathcal{A}^{-1}w_1, \mathcal{A}^{-1}w_2), \tag{26}$$

inducing the norm $|||w|||_* = a_{\mathrm{Re}}(\mathcal{A}^{-1}w, \mathcal{A}^{-1}w)^{1/2} = |||\mathcal{A}^{-1}w|||$ on V^*, equivalent to $\|w\|_*$.

The Lax–Milgram operator A is defined by restriction of $\mathcal{A}$ to an operator in H, i.e.,

$$Av = \mathcal{A}v \quad \text{for } v \in D(A) := \mathcal{A}^{-1}(H). \tag{27}$$

So $D(A)$ consists of the $u \in V$ for which some $f \in H$ fulfils $(f \,|\, v) = a(u,v)$ for all $v \in V$.

The reader may consult ([20], Section 12.4) for elementary proofs of the following: A is closed in H, with $D(A)$ dense in H as well as in V; in H also $\mathcal{A}'$ has these properties, and it equals the adjoint of A in H; i.e., $A'|_{\mathcal{A}'^{-1}(H)} = A^*$. As A is closed, $D(A)$ is a Hilbert space with the graph norm $\|v\|^2_{D(A)} = |v|^2 + |Av|^2$, and $D(A) \hookrightarrow V$ is bounded due to (22). Geometrically, $\sigma(A)$ and $\nu(A)$ are contained in the sector of $z \in \mathbb{C}$ given by

$$|\operatorname{Im} z| \leq C_3 C_4^{-1} \operatorname{Re} z. \tag{28}$$

Actually $0 \in \rho(A)$ since a is V-elliptic, so $A^{-1} \in \mathbb{B}(H)$; moreover $m(A) \geq C_1 C_4 / C_2 > 0$.

Both the closed operator A in H and its extension $\mathcal{A} \in \mathbb{B}(V, V^*)$ are used throughout. (For simplicity, they were both denoted by A in the introduction, though.)

2.2. The Self-Adjoint Case

As is well known, if A is selfadjoint, i.e., $A^* = A$ (or $a^* = a$), and has *compact* inverse, then H has an orthonormal basis of eigenvectors of A, which can be scaled to orthonormal bases of V and V^*. This is recalled, because our results can be given a more explicit form in this case, e.g., for $-\Delta$ in (1).

The properties that A is selfadjoint, closed, and densely defined with dense range in H carry over to A^{-1} (e.g., ([20], Theorem 12.7)), so when A^{-1} in addition is compact in H (e.g., if $V \hookrightarrow H$ is compact), then the spectral theorem for compact selfadjoint operators states that H has an orthonormal basis (e_j) consisting of eigenvectors of A^{-1}, where the eigenvalues μ_j of A^{-1} by the positivity can be ordered such that

$$\mu_1 \geq \mu_2 \geq \ldots \geq \mu_j \geq \cdots > 0, \quad \text{with } \mu_j \to 0 \text{ if } j \to \infty. \tag{29}$$

The orthonormal basis (e_j) also consists of eigenvectors of A with eigenvalues $\lambda_j = 1/\mu_j$. Hence $\sigma(A) = \sigma_{\text{point}}(A) = \{\, \lambda_j \mid j \in \mathbb{N} \,\}$. Indeed, $\sigma_{\text{res}}(A) = \emptyset$ as $A^* = A$; and $A^{-1} \in \mathbb{B}(H)$ while $A - \nu I = (\nu^{-1} I - A^{-1})\nu A$ has a bounded inverse for $\nu \neq \lambda_j$, as $\nu^{-1} \notin \sigma(A^{-1})$.

As $a_{\text{Re}} = a$ here, V is now renormed by $|\!|\!| v |\!|\!|^2 = a(v,v)$. However, if moreover V is considered with $a(u,v)$ as inner product, then $\mathcal{A}\colon V \to V^*$ is the Riesz isometry; and one has

Fact 1. *For every $v \in V$ the H-expansion $v = \sum_{j=1}^{\infty} (v \,|\, e_j) e_j$ converges in V. Moreover, the sequence $(e_j/\sqrt{\lambda_j})_{j\in\mathbb{N}}$ is an orthonormal basis for V, and $|\!|\!| v |\!|\!|^2 = \sum_{j=1}^{\infty} \lambda_j |(v \,|\, e_j)|^2$.*

Proof. The $e_j/\sqrt{\lambda_j}$ are orthonormal in V since $a(e_j, e_k) = \langle \mathcal{A} e_j, e_k \rangle = \lambda_j (e_j \,|\, e_k)$, cf. (25). They also yield a basis for V since similarly $w \in V \ominus \operatorname{span}(e_j/\sqrt{\lambda_j})$ implies $w = 0$. As $\lambda_j > 0$, expansion of any v in V gives

$$v = \sum_{j=1}^{\infty} a(v, \lambda_j^{-1/2} e_j) \lambda_j^{-1/2} e_j = \sum_{j=1}^{\infty} \overline{a(e_j, v)} \lambda_j^{-1} e_j = \sum_{j=1}^{\infty} (v \,|\, e_j) e_j, \tag{30}$$

whence the rightmost side converges in V. This means that $v = \sum_{j=1}^{\infty} \sqrt{\lambda_j} (v \,|\, e_j) e_j/\sqrt{\lambda_j}$ is an orthogonal expansion in V, whence $|\!|\!| v |\!|\!|^2$ has the stated expression. □

For V^* the set-up (26), (25) here gives $(w_1 \,|\, w_2)_{V^*} = a(\mathcal{A}^{-1} w_1, \mathcal{A}^{-1} w_2) = \langle w_1, \mathcal{A}^{-1} w_2 \rangle$.

Fact 2. *For every $w \in V^*$ the expansion $w = \sum_{j=1}^{\infty} \langle w, e_j \rangle e_j$ converges in V^*. Moreover, the sequence $(\sqrt{\lambda_j} e_j)_{j\in\mathbb{N}}$ is an orthonormal basis of V^* and $|\!|\!| w |\!|\!|_*^2 = \sum_{j=1}^{\infty} \lambda_j^{-1} |\langle w, e_j \rangle|^2$.*

Proof. $(\sqrt{\lambda_j} e_j)$ is orthonormal as $(e_j \,|\, e_k)_{V^*} = \langle e_j, \mathcal{A}^{-1} e_k \rangle = \lambda_k^{-1} (e_j \,|\, e_k)$; and if $w \in V^*$ for all j fulfils $0 = \langle e_j, \mathcal{A}^{-1} w \rangle = (e_j \,|\, \mathcal{A}^{-1} w)$, then $w = 0$ as $\mathcal{A}^{-1}$ is injective. Therefore

$$w = \sum_{j=1}^{\infty} (w \,|\, \lambda_j^{1/2} e_j)_{V^*} \lambda_j^{1/2} e_j = \sum_{j=1}^{\infty} \langle w, \mathcal{A}^{-1} e_j \rangle \lambda_j e_j = \sum_{j=1}^{\infty} \langle w, e_j \rangle e_j, \tag{31}$$

so the rightmost side converges in V^*, and the expression for $|\!|\!| w |\!|\!|_*^2$ results. □

2.3. Semigroups

Assuming that the reader is familiar with the theory of semigroups $e^{t\mathbf{A}}$, we review a few needed facts in a setting with a general complex Banach space B. The books of Pazy [23], Tanabe [9] and Yosida [19] may serve as general references.

The generator is $\mathbf{A}x = \lim_{t\to 0^+} \frac{1}{t}(e^{t\mathbf{A}}x - x)$, with domain $D(\mathbf{A})$ consisting of the $x \in B$ for which the limit exists. $\mathbf{A}$ is a densely defined, closed linear operator in B that for certain $\omega \geq 0$ and $M \geq 1$ satisfies $\|(\mathbf{A} - \lambda)^{-n}\|_{\mathbb{B}(B)} \leq M/(\lambda - \omega)^n$ for $\lambda > \omega$, $n \in \mathbb{N}$.

The corresponding semigroup of operators is written $e^{t\mathbf{A}}$, it belongs to $\mathbb{B}(B)$ with

$$\|e^{t\mathbf{A}}\|_{\mathbb{B}(B)} \leq Me^{\omega t} \quad \text{for } 0 \leq t < \infty. \tag{32}$$

Its basic properties are that $e^{t\mathbf{A}}e^{s\mathbf{A}} = e^{(s+t)\mathbf{A}}$ for $s, t \geq 0$, $e^{0\mathbf{A}} = I$, $\lim_{t\to 0^+} e^{t\mathbf{A}}x = x$ for $x \in B$, and the first of these gives at once the range inclusions

$$R(e^{(s+t)\mathbf{A}}) \subset R(e^{t\mathbf{A}}) \subset B. \tag{33}$$

The following well-known theorem gives a criterion for $\mathbf{A}$ to generate an analytic semigroup that is uniformly bounded, i.e., has $\omega = 0$. It summarises the most relevant parts of Theorems 1.7.7 and 2.5.2 in [23], and it involves sectors of the form

$$\Sigma := \left\{ \lambda \in \mathbb{C} \;\middle|\; |\arg \lambda| < \frac{\pi}{2} + \theta \right\} \cup \{0\}. \tag{34}$$

Theorem 3. *If $\theta \in]0, \frac{\pi}{2}[$ and $M > 0$ are such that the resolvent set $\rho(\mathbf{A}) \supseteq \Sigma$ and*

$$\|(\lambda I - \mathbf{A})^{-1}\|_{\mathbb{B}(B)} \leq \frac{M}{|\lambda|}, \quad \text{for } \lambda \in \Sigma,\ \lambda \neq 0, \tag{35}$$

then $\mathbf{A}$ generates an analytic semigroup $e^{z\mathbf{A}}$ for $|\arg z| < \theta$, for which $\|e^{z\mathbf{A}}\|$ is bounded for $|\arg z| \leq \theta' < \theta$, and $e^{t\mathbf{A}}$ is differentiable in $\mathbb{B}(B)$ for $t > 0$ with $(e^{t\mathbf{A}})' = \mathbf{A}e^{t\mathbf{A}}$. Here

$$\|\mathbf{A}e^{t\mathbf{A}}\|_{\mathbb{B}(B)} \leq \frac{c}{t} \quad \text{for } t > 0. \tag{36}$$

Furthermore, if $e^{t\mathbf{A}}$ is analytic, $u' = \mathbf{A}u$, $u(0) = u_0$ is uniquely solved by $u(t) = e^{t\mathbf{A}}u_0$ for *every* $u_0 \in B$.

2.3.1. Injectivity

Often it is crucial to know whether the semigroup $e^{t\mathbf{A}}$ consists of *injective* operators. Injectivity is, e.g., equivalent to the geometric property that the trajectories of two solutions $e^{t\mathbf{A}}v_0$ and $e^{t\mathbf{A}}w_0$ of $u' = \mathbf{A}u$ have no point of confluence in B for $v_0 \neq w_0$.

However, the literature seems to have focused on examples with non-invertibility of $e^{t\mathbf{A}}$, e.g., ([23], Example 2.2.1). However, injectivity always holds in the analytic case, as we now show:

Proposition 1. *When a semigroup $e^{z\mathbf{A}}$ on a complex Banach space B is analytic $S \to \mathbb{B}(B)$ in the sector $S = \{ z \in \mathbb{C} \mid |\arg z| < \theta \}$ for some $\theta > 0$, then $e^{z\mathbf{A}}$ is* injective *for all $z \in S$.*

Proof. Let $e^{z_0\mathbf{A}}u_0 = 0$ hold for some $u_0 \in B$, $z_0 \in S$. The analyticity of $e^{z\mathbf{A}}$ in S carries over to the map $f\colon z \mapsto e^{z\mathbf{A}}u_0$, and to $g_v\colon z \mapsto \langle v, f(z) \rangle$ for arbitrary v in the dual space B'. So g_v has in a ball $B(z_0, r) \subset S$ the Taylor expansion

$$g_v(z) = \sum_{n=0}^{\infty} \frac{1}{n!} \langle v, f^{(n)}(z_0) \rangle (z - z_0)^n. \tag{37}$$

By the properties of analytic semigroups (cf. ([23], Lemma 2.4.2)) and of u_0,

$$f^{(n)}(z_0) = \mathbf{A}^n e^{z_0\mathbf{A}}u_0 = 0 \quad \text{for all } n \geq 0, \tag{38}$$

so that $g_v \equiv 0$ holds on $B(z_0, r)$ and consequently on S by unique analytic extension.

Now $f(z_1) \neq 0$ would yield $g_v(z_1) \neq 0$ for a suitable v in B', hence $f \equiv 0$ on S and

$$u_0 = \lim_{t \to 0} e^{t\mathbf{A}} u_0 = \lim_{t \to 0} f(t) = 0, \tag{39}$$

since $e^{t\mathbf{A}}$ is a strongly continuous semigroup. Altogether $Z(e^{z_0\mathbf{A}}) = \{0\}$ is proved. □

Remark 1. *We have only been able to track a claim of the injectivity in Proposition 1 in case $z > 0$, $\theta \leq \pi/4$ and B is a Hilbert space; cf. Showalter's paper [17]. However, his proof is flawed, as $\mathbf{A}^2$ is non-accretive for some $\mathbf{A}$ with $\theta \leq \pi/4$, cf. the counter-example in Remark 9 below.*

Remark 2. *Injectivity also follows directly when $\mathbf{A}$ is defined on a Hilbert space H having an orthonormal basis $(e_n)_{n\in\mathbb{N}}$ such that $\mathbf{A}e_j = \lambda_j e_j$: Clearly $e^{t\mathbf{A}}e_j = e^{t\lambda_j}e_j$ as both sides satisfy $x' - \mathbf{A}x = 0$, $x(0) = e_j$. So if $e^{t\mathbf{A}}v = 0$, boundedness of $e^{t\mathbf{A}}$ gives*

$$0 = e^{t\mathbf{A}}v = \sum (v \,|\, e_j) e^{t\mathbf{A}} e_j = \sum (v \,|\, e_j) e^{t\lambda_j} e_j, \tag{40}$$

so that $v \perp e_j$ for all j, and thus $v \in \operatorname{span}(e_n)^\perp = H^\perp = \{0\}$. Hence $e^{t\mathbf{A}}$ is invertible for such $\mathbf{A}$.

We have chosen to use the symbol $e^{-t\mathbf{A}}$ to denote the inverse of the *analytic* semigroup $e^{t\mathbf{A}}$ generated by $\mathbf{A}$, consistent with the case in which $e^{t\mathbf{A}}$ does form a group in $\mathbb{B}(B)$, i.e.,

$$e^{-t\mathbf{A}} := (e^{t\mathbf{A}})^{-1} \qquad \text{for all } t \in \mathbb{R}. \tag{41}$$

This notation is convenient for our purposes (with some diligence).

For simplicity we observe the following when $B = H$ is a Hilbert space and $t > 0$: clearly $e^{-t\mathbf{A}}$ maps $D(e^{-t\mathbf{A}}) = R(e^{t\mathbf{A}})$ bijectively onto H, and it is an unbounded closed operator in H. As $(e^{t\mathbf{A}})^* = e^{t\mathbf{A}^*}$ also is analytic, so that $Z(e^{t\mathbf{A}^*}) = \{0\}$ by Proposition 1, we have $\overline{D(e^{-t\mathbf{A}})} = H$, i.e., the domain is dense in H.

A partial group phenomenon and other algebraic properties are collected here:

Proposition 2. *The inverses $e^{-t\mathbf{A}}$ in (41) form a semigroup of unbounded operators,*

$$e^{-t\mathbf{A}} e^{-s\mathbf{A}} = e^{-(t+s)\mathbf{A}} \qquad \textit{for } t, s \geq 0. \tag{42}$$

This extends to $(s, t) \in \,]-\infty, 0] \times \mathbb{R}$, but the right-hand side may be unbounded for $t + s > 0$.

Moreover, as unbounded operators the $e^{-t\mathbf{A}}$ commute with $e^{s\mathbf{A}} \in \mathbb{B}(H)$, i.e.,

$$e^{s\mathbf{A}} e^{-t\mathbf{A}} \subset e^{-t\mathbf{A}} e^{s\mathbf{A}} \qquad \textit{for } t, s \geq 0, \tag{43}$$

and there is a descending chain of domain inclusions

$$D(e^{-t'\mathbf{A}}) \subset D(e^{-t\mathbf{A}}) \subset H \qquad \textit{for } 0 < t < t'. \tag{44}$$

Proof. When $s, t \geq 0$, clearly $e^{-t\mathbf{A}} e^{-s\mathbf{A}} e^{(s+t)\mathbf{A}} = I_H$ holds, so that $e^{-(s+t)\mathbf{A}} \subset e^{-t\mathbf{A}} e^{-s\mathbf{A}}$; but equality necessarily holds, as the injection $e^{-t\mathbf{A}} e^{-s\mathbf{A}}$ cannot be a proper extension of the surjection $e^{-(s+t)\mathbf{A}}$. Whence (42). For $t + s \geq 0 \geq s$ this yields $e^{-t\mathbf{A}} e^{-s\mathbf{A}} = e^{-(t+s)\mathbf{A}} e^{s\mathbf{A}} e^{-s\mathbf{A}} = e^{-(t+s)\mathbf{A}}$. The case $-s > t \geq 0$ is similar.

Also the commutation follows at once, for the semigroup property gives

$$e^{s\mathbf{A}} e^{-t\mathbf{A}} = e^{-t\mathbf{A}} e^{t\mathbf{A}} e^{s\mathbf{A}} e^{-t\mathbf{A}} = e^{-t\mathbf{A}} e^{(s+t)\mathbf{A}} e^{-t\mathbf{A}} = e^{-t\mathbf{A}} e^{s\mathbf{A}} I_{R(e^{t\mathbf{A}})}, \tag{45}$$

where the right-hand side is a restriction of $e^{-t\mathbf{A}} e^{s\mathbf{A}}$. Finally (33) yields (44). □

Remark 3. *$D(e^{-t\mathbf{A}}e^{s\mathbf{A}}) = D(e^{-(t-s)\mathbf{A}})$ holds in (43), because (42) extends to negative s as stated. Hence (43) is a strict inclusion if the the first one in (44) is so for all t, t'.*

2.3.2. Some Regularity Properties

As a preparation we treat a few regularity questions for $s \mapsto e^{(t-s)\mathbf{A}}f(s)$, where the analytic operator function $E(s) = e^{(t-s)\mathbf{A}}$ has a singularity at $s = t$; cf. (36). This will be controlled when $f \in L_1(0,t;B)$.

That $Ef = e^{(t-\cdot)\mathbf{A}}f$ also is in $L_1(0,t;B)$ is undoubtedly known. So let us recall briefly how to prove it strongly measurable, i.e., to find a sequence of simple functions converging pointwise to $E(s)f(s)$ for a.e. $s \in [0,t]$; cf. [19]. Now f can be so approximated by a sequence (f_n), and E can by its continuity $[0,t[\to \mathbb{B}(B)$ also be approximated pointwise for $s < t$ by E_n defined on each subinterval $[t(j-1)2^{-n}, tj2^{-n}[$, $j = 1,\ldots,2^n$, as the value of E at the left end point. Then $Ef = \lim_n E_n f_n$ on $[0,t]$ a.e. Therefore $e^{(t-\cdot)\mathbf{A}}f \in L_1(0,t;B)$ follows directly from (32),

$$\|e^{(t-\cdot)\mathbf{A}}f\|_{L_1(0,t;B)} \le \int_0^t \|e^{(t-s)\mathbf{A}}\|\,\|f(s)\|\,ds \le Me^{\omega t}\|f\|_{L_1(0,t;B)}. \tag{46}$$

Moreover, $\langle \eta, e^{(t-\cdot)\mathbf{A}}f\rangle$ is seen to be in $L_1(0,t)$ by majorizing with $\|e^{(t-s)\mathbf{A}}f(s)\|_B\|\eta\|_{B^*}$, for strong measurability implies weak measurability; cf. ([24], Section IV.5 appendix).

The main concern is to obtain a Leibniz rule for the derivative:

$$\partial_s(e^{(T-s)\mathbf{A}}w(s)) = (-\mathbf{A})e^{(T-s)\mathbf{A}}w(s) + e^{(T-s)\mathbf{A}}\partial_s w(s). \tag{47}$$

For $w \in C^1(0,T;B)$ this is unproblematic for $s < T$: $w(s+h) = w(s) + h\partial_s w(s) + o(h)$, where $o(h)/h \to 0$ for $h \to 0$; and the operator is differentiable in $\mathbb{B}(B)$ for $s < T$, cf. Theorem 3, so that $e^{(T-(s+h))\mathbf{A}} = e^{(T-s)\mathbf{A}} + h(-\mathbf{A})e^{(T-s)\mathbf{A}} + o(h)$. Hence a multiplication of the two expansions gives the right-hand side of (47) to the first order in h. The Leibniz rule is more generally valid in the vector distribution sense:

Proposition 3. *When $\mathbf{A}$ generates an analytic semigroup on a Banach space B and $w \in H^1(0,T;B)$, then the Leibniz rule (47) holds in $\mathcal{D}'(0,T;B)$.*

Proof. It suffices to cover the case $\omega = 0$, for the other cases then follow by applying the formula to the semigroup $e^{-\omega t}e^{t\mathbf{A}}$ generated by $\mathbf{A} - \omega I$. For $w \in H^1(0,T;B)$ the standard convolution procedure gives a sequence (w_k) in $C^1([0,T];B)$ such that

$$w_k \to w \quad \text{in} \quad L_2(0,T;B), \qquad w_k' \to w' \quad \text{in} \quad L_{2,\text{loc}}(0,T;B). \tag{48}$$

For arbitrary $\phi \in C_0^\infty(\,]0,T[\,)$, we find using the Bochner inequality that

$$\|\int_0^T e^{(T-s)\mathbf{A}}(w(s) - w_k(s))\phi(s)\,ds\|_B \le C\|w(s) - w_k(s)\|_{L_2(0,T;B)}, \tag{49}$$

with $C = M(\int_{\operatorname{supp}\phi}|\phi(s)|^2\,ds)^{1/2}$, where M is the constant in (32).

Hence $e^{(T-s)\mathbf{A}}w_k \to e^{(T-s)\mathbf{A}}w$ in $\mathcal{D}'(0,T;B)$, so via the C^1-case above, as ∂_s is continuous in $\mathcal{D}'$, we get

$$\begin{aligned}\partial_s(e^{(T-s)\mathbf{A}}w) &= \lim_{k\to\infty}(\partial_s(e^{(T-s)\mathbf{A}}w_k)) \\ &= \lim_{k\to\infty}((-\mathbf{A})e^{(T-s)\mathbf{A}}w_k) + \lim_{k\to\infty}(e^{(T-s)\mathbf{A}}\partial_s w_k) = (-\mathbf{A})e^{(T-s)\mathbf{A}}w + e^{(T-s)\mathbf{A}}\partial_s w.\end{aligned} \tag{50}$$

Indeed, the last limits exist in $\mathcal{D}'(0,T;B)$ by the choice of w_k, for if $\epsilon > 0$ is small enough,

$$\| \int_{\operatorname{supp}\phi} e^{(T-s)\mathbf{A}}(w'(s) - w_k'(s))\phi(s)\,ds\|_B \le c\int_{\varepsilon}^{T-\varepsilon} \|w'(s) - w_k'(s)\|_B\,ds, \tag{51}$$

$$\| \int_0^T (-\mathbf{A})e^{(T-s)\mathbf{A}}(w(s) - w_k(s))\phi(s)\,ds\|_B \le \tilde{C}\|w - w_k\|_{L_2(0,T;B)} \tag{52}$$

with $\tilde{C} = (\int_{\operatorname{supp}\phi} |\frac{c\phi(s)}{T-s}|^2\,ds)^{1/2}$, using the bound on $(-\mathbf{A})e^{(T-s)\mathbf{A}}$ in Theorem 3. $\square$

3. Functional Analysis of Initial Value Problems

Having set the scene by recalling elliptic Lax–Milgram operators $\mathcal{A}$ in Gelfand triples (V, H, V^*) in Section 2.1, we now discuss solutions of the classical initial value problem

$$\begin{cases} \partial_t u + \mathcal{A}u = f & \text{in } \mathcal{D}'(0,T;V^*) \\ u(0) = u_0 & \text{in } H. \end{cases} \tag{53}$$

By definition of vector distributions, the above equation means that for every scalar test function $\varphi \in C_0^\infty(]0,T[)$ one has $\langle u, -\varphi' \rangle + \langle \mathcal{A}u, \varphi \rangle = \langle f, \varphi \rangle$ as an identity in V^*.

First we recall the fundamental theorem for vector functions from ([10], Lemma III.1.1). Further below, it will be crucial for obtaining a solution formula for (53).

Lemma 1. *For a Banach space B and $u, g \in L_1(a,b;B)$ the following are equivalent:*

(i) *u is a.e. equal to a primitive function of g, i.e., for some vector $\xi \in B$*

$$u(t) = \xi + \int_a^t g(s)\,ds \quad \text{for a.e. } t \in [a,b]. \tag{54}$$

(ii) *For each test function $\phi \in C_0^\infty(]a,b[)$ one has $\int_a^b u(t)\phi'(t)\,dt = -\int_a^b g(t)\phi(t)\,dt$.*

(iii) *For each η in the dual space B', $\frac{d}{dt}\langle \eta, u\rangle = \langle \eta, g\rangle$ holds in $\mathcal{D}'(a,b)$.*

In the affirmative case, $u' = g$ as vector distributions in $\mathcal{D}'(a,b;B)$ by (ii), the right-hand side in (i) is a continuous representative of u such that $\xi = u(a)$ and

$$\sup_{a\le t\le b} \|u(t)\|_B \le (b-a)^{-1}\|u\|_{L_1(a,b;B)} + \|g\|_{L_1(a,b;B)}. \tag{55}$$

Remark 4. *Lemma 1 is proved in [10], except for the estimate (55): the continuous function $\|u(t)\|_B$ attains its minimum at some $t_0 \in [a,b]$, so applying the Bochner inequality in (i) and the Mean Value Theorem,*

$$\|u(t)\|_B \le \|u(t_0)\|_B + |\int_{t_0}^t \|g(t)\|_B\,dt| \le \frac{1}{b-a}\int_a^b \|u(t)\|_B\,dt + \int_a^b \|g(t)\|_B\,dt. \tag{56}$$

This yields (55), hence the Sobolev embedding $W^{1,1}(a,b;B) \hookrightarrow C([a,b];B)$. If furthermore $u, g \in L_2(a,b;B)$, we get the Sobolev embedding $H^1(a,b;B) \hookrightarrow C([a,b];B)$ similarly,

$$\sup_{a\le t\le b} \|u(t)\|_B \le (b-a)^{-1/2}\|u\|_{L_2(a,b;B)} + (b-a)^{1/2}\|g\|_{L_2(a,b;B)} \le c\|u\|_{H^1(a,b;B)}. \tag{57}$$

Secondly we recall the Leibniz rule $\frac{d}{dt}(f(t)\,|\,g(t)) = (f'(t)\,|\,g(t)) + (f(t)\,|\,g'(t))$ valid for $f, g \in C^1([0,T];H)$. The well-known generalization below was proved in real vector spaces in ([10], Lemma III.1.2) for $u = v$. We briefly extend this to the general complex case, which we mainly use to obtain that $\partial_t|u|^2 = 2\operatorname{Re}\langle u', u \rangle$, though also $u \ne v$ will be needed.

Lemma 2. *If $u, v \in L_2(0,T;V) \cap H^1(0,T;V^*)$, then $t \mapsto (u(t)\,|\,v(t))$ is in $L_1(0,T)$ and*

$$\frac{d}{dt}(u\,|\,v) = \langle u', v\rangle + \overline{\langle v', u\rangle} \quad \text{in } \mathcal{D}'(\prime, \mathcal{T}). \tag{58}$$

Furthermore, u and v have continuous representatives on $[0,T]$, i.e., $u, v \in C([0,T];H)$.

Proof. Let $u, v \in L_2(0,T;V)$ with distributional derivatives $u', v' \in L_2(0,T;V^*)$. As in the proof of Proposition 3 we obtain $u_m \in C^\infty([0,T];V)$ such that $u_m \to u$ in $L_2(0,T;V)$ while $u'_m \to u'$ in $L_{2,\mathrm{loc}}(0,T;V^*)$. Similarly v gives rise to v_m.

By continuity of inner products, the function $(u\,|\,v)$ is measurable on $[0,T]$ for $u, v \in L_2(0,T;V)$, and $\int_0^T |(u\,|\,v)|\,dt < \infty$. Sesquilinearity yields $(u_m\,|\,v_m) \to (u\,|\,v)$ in $L_1(0,T)$ for $m \to \infty$, while both $\langle u'_m, v_m\rangle \to \langle u', v\rangle$ and $\langle v'_m, u_m\rangle \to \langle v', u\rangle$ hold in $L_{2,\mathrm{loc}}(0,T)$, hence in $\mathcal{D}'(0,T)$.

As differentiation is continuous in $\mathcal{D}'(0,T)$, one finds from the C^1-case and (21) that

$$\frac{d}{dt}(u\,|\,v) = \lim_m \frac{d}{dt}(u_m\,|\,v_m) = \lim_m (u'_m\,|\,v_m) + \lim_m \overline{(v'_m\,|\,u_m)} = \langle u', v\rangle + \overline{\langle v', u\rangle}. \tag{59}$$

Taking $v = u$ the function $t \mapsto |u(t)|^2$ is seen to be in $W^{1,1}(0,T) \subset C([0,T])$, and since any $u \in H^1(0,T;V^*)$ is continuous in V^* by Remark 4, one can also here obtain from Lemma III.1.4 in [10] that $u\colon [0,T] \to H$ is continuous. Similarly for v. □

3.1. Existence and Uniqueness

In our presentation the following result is a cornerstone, relying on the full framework in Section 2.1; in particular A need not be selfadjoint:

Theorem 4. *Let V be a separable Hilbert space with $V \subseteq H$ algebraically, topologically and densely, cf. (19) and (20), and let $\mathcal{A}\colon V \to V^*$ be the bounded Lax–Milgram operator induced by a V-elliptic sesquilinear form, cf. (25). When $u_0 \in H$ and $f \in L_2(0,T;V^*)$ are given, then (53) has a uniquely determined solution $u(t)$ belonging to the space*

$$X = L_2(0,T;V) \bigcap C([0,T];H) \bigcap H^1(0,T;V^*). \tag{60}$$

We omit a proof of this theorem, as it is a special case of a more general result of Lions and Magenes ([8], Section 3.4.4) on t-dependent forms $a(t;u,v)$. Clearly the conjunction of $u \in L_2(0,T;V)$ and $u' \in L_2(0,T;V^*)$, which appears in [8], is equivalent to the claim in (60) that u belongs to the intersection of $L_2(0,T,V)$ and $H^1(0,T;V^*)$.

Alternatively one can use Theorem III.1.1 in Temam's book [10], where proof is given using Lemma 1 to reduce to the scalar differential equation $\partial_t\langle u, \eta\rangle + a(u,\eta) = \langle f, \eta\rangle$ in $\mathcal{D}'(0,T)$, for $\eta \in V$, which is treated by Faedo–Galerkin approximation and basic functional analysis. His proof extends straightforwardly, from a specific triple (H, V, a) for the Navier-Stokes equations, to the general set-up in Section 2.1, also when $A^* \neq A$.

However, either way, we need the finer theory described in the next two subsections.

3.2. Well-Posedness

We now substantiate that the unique solution from Theorem 4 depends continuously on the data, so that (53) is well-posed in the sense of Hadamard. First we note that the solution in Theorem 4 is an element of the space X in (60), which is a Banach space when normed, as done throughout, by

$$\|u\|_X = \Big(\|u\|^2_{L_2(0,T;V)} + \sup_{0\le t\le T} |u(t)|^2 + \|u\|^2_{H^1(0,T;V^*)}\Big)^{1/2}. \tag{61}$$

To clarify a redundancy in this choice, we need a Sobolev inequality for vector functions.

Lemma 3. *There is an inclusion $L_2(0,T;V) \cap H^1(0,T;V^*) \subset C([0,T];H)$ and*

$$\sup_{0 \le t \le T} |u(t)|^2 \le (1 + \frac{C_2^2}{C_1^2 T}) \int_0^T \|u\|^2 \, dt + \int_0^T \|u'\|_*^2 \, dt. \tag{62}$$

Proof. If u belongs to the intersection, the continuity follows from Lemma 2, where the formula gives $\partial_t |u|^2 = 2 \operatorname{Re}\langle u', u \rangle$. By Lemma 1, integration of both sides entails

$$|u(t)|^2 \le |u(t_0)|^2 + \int_0^T (\|u\|^2 + \|u'\|_*^2) \, dt, \tag{63}$$

which by use of the Mean Value Theorem as in Remark 4 leads to the estimate. □

Remark 5. *In our solution set X in (60) one can safely omit the space $C([0,T];H)$, according to Lemma 3. Likewise $\sup |u|$ can be removed from $\|\cdot\|_X$, as one just obtains an equivalent norm (similarly for the term $\int_0^T \|u(t)\|_*^2 \, dt$ in (7)). Thus X is more precisely a Hilbertable space; we omit this detail in the sequel for the sake of simplicity. However, we shall keep X as stated in order to emphasize the properties of the solutions.*

The next result on stability is well known among experts, and while it may be derived from the abstract proofs in [8], we shall give a direct proof based on explicit estimates:

Corollary 1. *The unique solution u of (53), given by Theorem 4, depends continuously as an element of X on the data $(f, u_0) \in L_2(0,T;V^*) \oplus H$, i.e.,*

$$\|u\|_X^2 \le c(|u_0|^2 + \|f\|_{L_2(0,T;V^*)}^2). \tag{64}$$

That is, the solution operator $(f, u_0) \mapsto u$ is a bounded linear map $L_2(0,T;V^*) \oplus H \to X$.

Proof. Clearly $u \in L_2(0,T;V)$ while the functions u', f and $\mathcal{A}u$ belong to $L_2(0,T;V^*)$, so as an identity of integrable functions,

$$\operatorname{Re}\langle \partial_t u, u \rangle + \operatorname{Re}\langle \mathcal{A}u, u \rangle = \operatorname{Re}\langle f, u \rangle. \tag{65}$$

Hence Lemma 2 and the V-ellipticity gives

$$\partial_t |u|^2 + 2C_4 \|u\|^2 \le 2|\langle f, u \rangle| \le C_4^{-1} \|f\|_*^2 + C_4 \|u\|^2. \tag{66}$$

Using again that $|u(t)|^2$ and $\partial_t |u(t)|^2$ are in $L_1(0,T)$, taking $B = \mathbb{C}$ in Lemma 1 yields

$$|u(t)|^2 + C_4 \int_0^t \|u(s)\|^2 \, ds \le |u_0|^2 + C_4^{-1} \|f\|_{L_2(0,T;V^*)}^2. \tag{67}$$

For the first two contributions to the X-norm this gives

$$\sup_{0 \le t \le T} |u(t)|^2 \le |u_0|^2 + C_4^{-1} \|f\|_{L_2(0,T;V^*)}^2, \tag{68}$$

$$\|u\|_{L_2(0,T;V)}^2 \le C_4^{-1} |u_0|^2 + C_4^{-2} \|f\|_{L_2(0,T;V^*)}^2. \tag{69}$$

Since u solves (53) it is clear that $\|\partial_t u(t)\|_*^2 \le (\|f(t)\|_* + \|\mathcal{A}u\|_*)^2$, so we get

$$\int_0^T \|\partial_t u(t)\|_*^2 \, dt \le 2 \int_0^T \|f(t)\|_*^2 \, dt + 2\|\mathcal{A}\|_{\mathbb{B}(V,V^*)}^2 \int_0^T \|u\|^2 \, dt, \tag{70}$$

which upon substitution of (69) altogether shows (64). □

3.3. The First Order Solution Formula

We now supplement the well-posedness by a direct proof of the variation of constants formula, which requires that the extended Lax–Milgram operator $\mathcal{A}$ generates an analytic semigroup in V^*. This is known, cf. [9], but lacking a concise proof in the literature, we begin by analysing A in H:

Lemma 4. *For a V-elliptic Lax–Milgram operator A, both $-A$ and $-A^*$ have the sector Σ in (34) in their resolvent sets for $\theta = \operatorname{arccot}(C_3/C_4)$ and they generate analytic semigroups on H. This holds verbatim for the extensions $-\mathcal{A}$ and $-\mathcal{A}'$ in V^*.*

Proof. To apply Theorem 3, we let $\lambda \neq 0$ be given in the sector Σ for some angle θ satisfying $0 < \theta < \operatorname{arccot}(C_3/C_4)$. Then it is clear that $\delta = -\operatorname{sgn}(\operatorname{Im}\lambda)\theta$ or $\delta = 0$ gives

$$\operatorname{Re}(e^{i\delta}\lambda) \geq 0. \tag{71}$$

In case $\delta \in \{\pm\theta\}$ a multiplication of the inequalities (28) by $-\sin\delta$ yields

$$-\sin\delta \operatorname{Im} a(u,u) \geq -C_3 C_4^{-1} \sin\theta \operatorname{Re} a(u,u). \tag{72}$$

In addition $C_\theta := C_4\cos\theta - C_3\sin\theta > 0$, because $\cot\theta > C_3 C_4^{-1}$. So for $u \in D(A)$,

$$\begin{aligned}
\operatorname{Re}(e^{i\delta}(a(u,u) + \lambda(u\,|\,u))) \geq \operatorname{Re}(e^{i\delta}a(u,u)) &= \cos\delta \operatorname{Re} a(u,u) - \sin\delta \operatorname{Im} a(u,u) \\
&\geq (\cos\theta - C_3 C_4^{-1}\sin\theta) \operatorname{Re} a(u,u) \\
&\geq C_\theta \|u\|^2.
\end{aligned} \tag{73}$$

This V-ellipticity holds also if $\delta = 0$, cf. (71), so $e^{i\delta}(A + \lambda I)$ is in any case bijective; and so is $-A - \lambda I$.

To bound $-(A + \lambda I)^{-1}$, we see from (73) that for $u \in D(A)$,

$$\begin{aligned}
|\lambda|(u\,|\,u) &\leq |((A+\lambda)u\,|\,u)| + |a(u,u)| \leq |((A+\lambda)u\,|\,u)| + C_3\|u\|^2 \\
&\leq (1 + C_3 C_\theta^{-1})|((A+\lambda)u\,|\,u)|.
\end{aligned} \tag{74}$$

This implies (35) for $-A$. Since A^* is the Lax–Milgram operator associated to the elliptic form a^*, the above also entails the statement for $-A^*$.

For $\mathcal{A}$ it follows at once from (73) that $\operatorname{Re}\langle e^{i\delta}(\mathcal{A}+\lambda)u, u\rangle \geq C_\theta\|u\|^2$ for $u \in V$. Hence $R(\mathcal{A} + \lambda I)$ is closed in V^*, and it is also dense since $R(\mathcal{A} + \lambda I) \supset R(A + \lambda I) = H$ by the above; i.e., $\mathcal{A} + \lambda I$ is surjective. Mimicking (74), we get for $u \neq 0$, $\|w\| = 1$, both in V,

$$|\lambda| \cdot \|u\|_* \leq \sup_w |\langle(\mathcal{A}+\lambda)u, w\rangle| + C_3 C_\theta^{-1} |\langle(\mathcal{A}+\lambda)u, \frac{1}{\|u\|}u\rangle| \leq c\|(\mathcal{A}+\lambda)u\|_*. \tag{75}$$

This yields injectivity of $\mathcal{A} + \lambda I$ and the resolvent estimate. $\mathcal{A}'$ is covered through a^*. □

We denote by $e^{-t\mathcal{A}}$ the semigroup generated by $-\mathcal{A}$ on V^*, to distinguish it from e^{-tA} on H. Analogously for $e^{-t\mathcal{A}'} \in \mathbb{B}(V^*)$. As $A \subset \mathcal{A}$ implies that $(\mathcal{A} + \lambda I)^{-1}|_H = (A + \lambda I)^{-1}$, and since A and $\mathcal{A}$ have the same sector Σ by Lemma 4, the well-known Laplace transformation formula, cf. ([23], Theorem 1.7.7), yields the corresponding fact, say $e^{-t\mathcal{A}}|_H = e^{-tA}$ for the semigroups:

Lemma 5. *For all $x \in H$ one has $e^{-t\mathcal{A}}x = e^{-tA}x$ as well as $e^{-t\mathcal{A}'}x = e^{-tA^*}x$.*

We could add that A and A^* are dissipative, as $m(A) > 0$, $m(A^*) > 0$ in H, so e^{-tA}, e^{-tA^*} are contractions for $t \geq 0$ by the Lumer–Philips theorem; cf. ([20], Corollary 14.12).

Using Lemmas 4 and 5, the announced formula results as an addendum to Theorem 4:

Theorem 5. *The unique solution u in X provided by Theorem 4 satisfies that*

$$u(t) = e^{-tA}u_0 + \int_0^t e^{-(t-s)\mathcal{A}} f(s)\, ds \qquad \text{for } 0 \le t \le T, \tag{76}$$

where each of the three terms belongs to X.

Proof. Once (76) has been shown, Theorem 4 applies in particular to cases with $f = 0$, yielding that $u(t)$ and hence $e^{-tA}u_0$ belongs to X. For general data (f, u_0) this means that the last term containing f necessarily is a member of X too.

To derive Formula (76) in the present general context, one should note that all terms in the equation $\partial_t u + \mathcal{A}u = f$ belong to the space $L_2(0, T; V^*)$. Therefore the operator $e^{-(T-t)\mathcal{A}}$ applies to both sides as an integration factor, yielding

$$e^{-(T-t)\mathcal{A}}\partial_t u(t) + e^{-(T-t)\mathcal{A}}\mathcal{A}u(t) = e^{-(T-t)\mathcal{A}} f(t). \tag{77}$$

Now $e^{-(T-t)\mathcal{A}}u(t)$ is in $L_1(0, T; V^*)$, cf. the argument prior to (46). For its derivative in $\mathcal{D}'(0, T; V^*)$ the Leibniz rule in Proposition 3 gives, as $u(t) \in V = D(\mathcal{A})$ for t a.e.,

$$\partial_t(e^{-(T-t)\mathcal{A}}u(t)) = e^{-(T-t)\mathcal{A}}\partial_t u(t) + e^{-(T-t)\mathcal{A}}\mathcal{A}u(t). \tag{78}$$

As both terms on the right-hand side are in $L_2(0, T; V^*)$, the implication (ii) $\Longrightarrow$ (i) in Lemma 1 gives

$$e^{-(T-t)\mathcal{A}}u(t) = e^{-T\mathcal{A}}u_0 + \int_0^t e^{-(T-s)\mathcal{A}} f(s)\, ds. \tag{79}$$

From this identity in $C([0, T]; V^*)$ formula (76) results in case $t = T$ by evaluation, when also Lemma 5 is used for the term containing u_0. However, obviously the above argument applies to any subinterval $[0, T_1] \subset [0, T]$, whence (76) is valid for all t in $[0, T]$. □

Alternatively one could conclude by applying $e^{-(T-s)\mathcal{A}} = e^{-(T-t)\mathcal{A}}e^{-(t-s)\mathcal{A}}$ in (79) and use the Bochner identity to commute $e^{-(T-t)\mathcal{A}}$ with the integral: as analytic semigroups like $e^{-(T-t)\mathcal{A}}$ are always injective, cf. Proposition 1, formula (76) then results at once.

For later reference we show similarly the next inequality:

Corollary 2. *The solution $e^{-t\mathcal{A}}u_0$ to the problem with $f = 0$ in Theorem 4 belongs to $L_2(0, T; V)$ and fulfils, for every $u_0 \in H$,*

$$\sup_{0 \le t \le T} (T - t)|e^{-t\mathcal{A}}u_0|^2 \le C_5 \int_0^T \|e^{-t\mathcal{A}}u_0\|^2\, dt. \tag{80}$$

Proof. It is seen from Theorem 5 that $u(t) = e^{-t\mathcal{A}}u_0$ always is in $L_2(0, T; V)$, as a member of X. By taking scalar products with $(T - \cdot)u$ on both sides of the differential equation, one obtains in $L_1(0, T)$ the identity

$$(T - t)\langle\, u'(t), u(t)\,\rangle + (T - t)a(u(t), u(t)) = 0. \tag{81}$$

Taking real parts here, applying Lemma 2 to u and integrating partially on $[t, T]$, one obtains

$$\int_t^T |u(s)|^2\, ds - (T - t)|u(t)|^2 = -2\int_t^T (T - s)\operatorname{Re} a(u(s), u(s))\, ds. \tag{82}$$

By reorganising this, a crude estimate yields the result at once for $C_5 = \frac{C_2}{C_1} + 2TC_3$. □

3.4. Non-Selfadjoint Dynamics

It is classical that $e^{-tA}u_0$ in (76) is a term that decays exponentially for $t \to \infty$ if A is self-adjoint and has compact inverse on H. This follows from the eigenfunction expansions, cf. the formulas in the introduction and Section 2.2, which imply for the 'height' function $h(t) = |e^{-tA}u_0|$ that $h(t) = \mathcal{O}(e^{-t \operatorname{Re} \lambda_1})$.

However, it is a much more precise dynamical property that $h(t)$ is a *strictly convex* function for $u_0 \neq 0$ (we refer to [25] for a lucid account of convex functions). Strict convexity is established below for wide classes of non-self-adjoint A, namely if A is hyponormal or such that A^2 is accretive.

Moreover, it seems to be a novelty that the *injectivity* of e^{-tA} provided by Proposition 1 implies the strict convexity. For simplicity we first explain this for the square $h(t)^2$.

Indeed, differentiating twice for $t > 0$ one finds for $u = e^{-tA}u_0$,

$$(h^2)'' = (-2\operatorname{Re}(Ae^{-tA}u_0 \mid e^{-tA}u_0))' = 2\operatorname{Re}(A^2u \mid u) + 2(Au \mid Au). \tag{83}$$

In case A^2 is accretive, that is when $m(A^2) \geq 0$, we may keep only the last term in (83) to get that $(h^2)''(t) \geq 2|Ae^{-tA}u_0|^2$, which for $u_0 \neq 0$ implies $(h^2)'' > 0$ as both A and e^{-tA} are injective; cf. (34) and Proposition 1. Hence h^2 is strictly convex for $t > 0$ if $m(A^2) \geq 0$.

Another case is when A is *hyponormal*. For an unbounded operator A this means, cf. the work of Janas [18], that

$$D(A) \subset D(A^*) \quad \text{with} \quad |A^*u| \leq |Au| \quad \text{for all } u \in D(A). \tag{84}$$

Note that if both A, A^* are hyponormal, then A is normal. This is a quite general class, but it fits most naturally into the present discussion:

For hyponormal A we have $R(e^{-tA}) \subset D(A) \subset D(A^*)$, which shows that $A^*e^{-tA}u_0$ is defined. Using this and hyponormality once more in (83), we get

$$(h^2)''(t) \geq (Au \mid A^*u) + (A^*u \mid Au) + |Au|^2 + |A^*u|^2 = |(A + A^*)e^{-tA}u_0|^2. \tag{85}$$

Now $(h^2)'' > 0$ follows for $u_0 \neq 0$ from injectivity of e^{-tA} and of $A + A^*$; the latter holds since $2a_{\operatorname{Re}}$ is V-elliptic. So h^2 is also strictly convex for hyponormal A.

Also on the closed half-line with $t \geq 0$ there is a result on non-selfadjoint dynamics. Here we return to $h(t)$ itself and normalise, at no cost, to $|u_0| = 1$ to get cleaner statements:

Proposition 4. *Let A denote a V-elliptic Lax–Milgram operator, defined from a triple (H, V, a), such that A is hyponormal, as above, or such that A^2 is accretive, and let u be the solution from Theorem 4 for $f = 0$ and $|u_0| = 1$. Then $h(t) = |u(t)|$ is strictly decreasing and* strictly convex *for $t \geq 0$ and differentiable from the right at $t = 0$ with*

$$h'(0) = -\operatorname{Re}(Au_0 \mid u_0) \qquad \text{for } u_0 \in D(A), \tag{86}$$

and generally

$$h'(0) \leq -m(A). \tag{87}$$

Remark 6. *The derivative $h'(0)$ might be $-\infty$ if $u_0 \in H \setminus D(A)$.*

Proof. By the convexity shown above, $(h^2)'$ is increasing. Since $m(A) > 0$ holds by the V-ellipticity, h^2 is strictly decreasing (and so is h) for $t > 0$ as

$$(h^2)'(t) = -2\operatorname{Re}(Ae^{-tA}u_0 \mid e^{-tA}u_0) \leq -2m(A)|e^{-tA}u_0|^2 < 0. \tag{88}$$

These properties give that $h' = (h^2)'/(2\sqrt{h^2})$ is *strictly* increasing for $t > 0$, so the Mean Value Theorem yields that $(h(t) - h(s))/(t - s) < (h(u) - h(t))/(u - t)$ for $0 < s < t < u$; i.e., h is strictly convex on $]0, \infty[$.

The inequality $h((1-\theta)t+\theta s) \leq (1-\theta)h(t)+\theta h(s)$, $\theta \in]0,1[$ now extends by continuity to $t=0$. So does strict convexity of h, using twice that the slope function is increasing.

By convexity h' is increasing for $t>0$, so $\lim_{t\to 0^+} h'(t) = \inf h' \geq -\infty$. For each $0<s<1$ the continuity of h yields $|e^{-tA}u_0| \geq s|u_0| = s$ for all sufficiently small $t \geq 0$. By the above formulas for h' and $(h^2)'$ we have $h'(t) = -\operatorname{Re}(Ae^{-tA}u_0 \,|\, e^{-tA}u_0)/|e^{-tA}u_0|$, so the Mean Value Theorem gives for some $\tau \in]0,t[$,

$$t^{-1}(h(t)-h(0)) = h'(\tau) \leq -m(A)s < 0. \tag{89}$$

Hence $h(0) > h(t)$ for all $t>0$. Moreover, the limit of $h'(\tau)$ was shown above to exist for $\tau \to 0^+$, so $h'(0)$ exists in $[-\infty, -m(A)]$. If $u_0 \in D(A)$ we may commute A with the semigroup in the formula for $h'(\tau)$, which by continuity gives $h'(0) = -\operatorname{Re}(Au_0 \,|\, u_0)$. □

Proposition 4 is a stiffness result for $u = e^{-tA}u_0$, due to strict convexity of $|e^{-tA}u_0|$. It is noteworthy that when $A \neq A^*$, then Proposition 4 gives conditions under which the eigenvalues in $\mathbb{C}\setminus\mathbb{R}$ (if any) never lead to oscillations in the *size* of the solution.

Remark 7. *Since $h'(0)$ is estimated in terms of the lower bound $m(A)$, it is the numerical range $\nu(A)$, rather than $\sigma(A)$, that controls* short-time *decay of the solutions $e^{-tA}u_0$.*

Remark 8. *In Proposition 4 we note that when A^2 is accretive, i.e., $m(A^2) \geq 0$, then A is* necessarily *sectorial with half-angle $\pi/4$; that is $\nu(A) \subset \{ z \in \mathbb{C} \mid |\arg(z)| \leq \pi/4 \}$. This may be seen as in ([17], Lemma 3), where reduction to bounded operators was made in order to invoke the operator monotonicity of the square root.*

Remark 9. *We take the opportunity to point out an error in ([17], Lemma 3), where it incorrectly was claimed that having half-angle $\pi/4$ also is sufficient for $m(A^2) \geq 0$. A counter-example is available already for A in $\mathbb{B}(H)$ (if $\dim H \geq 2$), as $A = X + \mathrm{i}\,Y$ for self-adjoint $X, Y \in \mathbb{B}(H)$: here $m(A) \geq 0$ if and only if $X \geq 0$, and we can even arrange that A has half-angle $\pi/4$, that is $|\operatorname{Im}(Av \,|\, v)| \leq \operatorname{Re}(Av \,|\, v)$ or $|(Yv \,|\, v)| \leq (Xv \,|\, v)$, by designing Y so that $-X \leq Y \leq X$. Here we may take $Y = \delta X + \lambda_1 U$, where $\delta > 0$ is small enough and U is a partial isometry that interchanges two eigenvectors v_1, v_2 of X with eigenvalues $\lambda_2 > \lambda_1 > 0$, $U = 0$ on $H \ominus \operatorname{span}(v_1, v_2)$. In fact, writing $v = c_1 v_1 + c_2 v_2 + v_\perp$ for $v_\perp \in H \ominus \operatorname{span}(v_1, v_2)$, since $v_1 \perp v_2$, the above inequalities for Y are equivalent to $2\lambda_1 |\operatorname{Re}(c_1 \bar{c}_2)| \leq \lambda_1(1-\delta)|c_1|^2 + (1-\delta)\lambda_2 |c_2|^2 + (1-\delta)(Xv_\perp \,|\, v_\perp)$, which by the positivity of X and Young's inequality is implied by $1/(1-\delta) \leq (1-\delta)\frac{\lambda_2}{\lambda_1}$, that is if $0 < \delta \leq 1 - \sqrt{\lambda_1/\lambda_2}$. Now, $m(A^2) \geq 0$ if and only if $|Xv|^2 \geq |Yv|^2$ for all v in H, but this will* always *be violated, as one can see from $|Yv|^2 = \delta^2 |Xv|^2 + \lambda_1^2 |Uv|^2 + 2\delta\lambda_1 \operatorname{Re}(Xv \,|\, Uv)$ by inserting $v = v_1$, for the last term drops out as $v_1 \perp v_2 = Uv_1$, so that actually $|Yv_1|^2 = (\delta^2+1)\lambda_1^2 > |Xv_1|^2$. Thus $A = \left(\begin{smallmatrix} \lambda & 0 \\ 0 & 4\lambda \end{smallmatrix}\right) + i\lambda \left(\begin{smallmatrix} \delta & 1 \\ 1 & 4\delta \end{smallmatrix}\right)$ is a counter-example in $\mathbb{C}^2$ for any $\lambda > 0$, $0 < \delta \leq 1/2$.*

Remark 10. *It is perhaps useful to emphasize the benefit from joining the two methods. Within semigroup theory the "mild solution" given in (76) is the only possible solution to (53); but as our class of solutions is larger, the extension of the old uniqueness argument in Theorem 5 was needed. Existence of a solution is for analytic semigroups classical if $f\colon [0,T] \to H$ is Hölder continuous, cf. ([23], Corollary 4.3.3). Using functional analysis, this gap to the weaker condition $f \in L_2(0,T;V^*)$ is bridged by Theorem 5, which states that the mild solution is indeed the solution in the space of vector distributions in Theorem 4; albeit at the expense that the generator A is a V-elliptic Lax–Milgram operator.*

4. Abstract Final Value Problems

In this section, we show for a Lax–Milgram operator $\mathcal{A}$ that the final value problem

$$\begin{cases} \partial_t u + \mathcal{A}u = f & \text{in } \mathcal{D}'(0,T;V^*), \\ \quad u(T) = u_T & \text{in } H, \end{cases} \tag{90}$$

is *well-posed* when the final data belong to an appropriate space, to be identified below. This is obtained via comparison with the initial value problem treated in Section 3.

4.1. A Bijection From Initial to Terminal States

According to Theorem 4, the solutions to the differential equation $u' + \mathcal{A}u = f$ are for fixed f parametrised by the initial states $u(0) \in H$. To study the terminal states $u(T)$ we note that (76) yields

$$u(T) = e^{-T\mathcal{A}}u(0) + \int_0^T e^{-(T-s)\mathcal{A}} f(s)\,ds. \tag{91}$$

This representation of $u(T)$ is essential in what follows, as it gives a bijective correspondence $u(0) \leftrightarrow u(T)$ between the initial and terminal states, as accounted for below.

First we analyse the integral term above by introducing the yield map $f \mapsto y_f$ given by

$$y_f = \int_0^T e^{-(T-s)\mathcal{A}} f(s)\,ds, \qquad f \in L_2(0,T;V^*). \tag{92}$$

Clearly y_f is a vector in V^* by definition of the integral (and since $C([0,T];V^*) \subset L_1(0,T;V^*)$). But actually it is in the smaller space H, for $y_f = u(T)$ holds in H when u is the solution for $u_0 = 0$ of (53), and then Corollary 1 yields an estimate of $\sup_{t\in[0,T]} |u(t)|$ by the L_2-norm of f; cf. (61). In particular, we have

$$|y_f| \leq c\|f\|_{L_2(0,T;V^*)}. \tag{93}$$

Moreover, $f \mapsto y_f$ is by (93) bounded $L_2(0,T;V^*) \to H$, and it has dense range in H containing all $x \in D(e^{\varepsilon A})$ for every $\varepsilon > 0$, for if in (92) we insert the piecewise continuous function

$$f_\varepsilon(s) = \mathbf{1}_{[T-\varepsilon,T]}(s)e^{(T-\varepsilon-s)A}(\frac{1}{\varepsilon}e^{\varepsilon A}x), \tag{94}$$

then the semigroup property gives $y_{f_\varepsilon} = \int_{T-\varepsilon}^T e^{-\varepsilon A}(\frac{1}{\varepsilon}e^{\varepsilon A}x)\,ds = \frac{1}{\varepsilon}\int_{T-\varepsilon}^T x\,ds = x$. However, standard operator theory gives the optimal result, that is, surjectivity:

Proposition 5. *The yield map $f \mapsto y_f$ is in $\mathbb{B}(L_2(0,T;V^*),H)$ and it is surjective, $R(y_f) = H$. Its adjoint in $\mathbb{B}(H, L_2(0,T;V))$ is the orbit map given by $v \mapsto e^{-(T-\cdot)A^*}v$.*

Proof. To determine the adjoint of $f \mapsto y_f$, we first calculate for $f \in L_2(0,T;H)$ so that the integrand in (92) belongs to $C([0,T];H)$. For $v \in H$ we get, using the Bochner identity twice,

$$(y_f \,|\, v) = \int_0^T (e^{-(T-s)A}f(s) \,|\, v)\,ds = \int_0^T (f(s) \,|\, e^{-(T-s)A^*}v)\,ds = \langle\, f, e^{-(T-s)A^*}v\,\rangle. \tag{95}$$

The last scalar product makes sense because $s \mapsto e^{-(T-s)A^*}v$ is in $L_2(0,T;V)$, as seen by applying Corollary 2 to the Lax–Milgram operator A^*, and $L_2(0,T;V)$ is the dual space to $L_2(0,T;V^*)$; cf. Remark 11 below. Since $L_2(0,T;H)$ is dense in $L_2(0,T;V^*)$, it follows by closure that the left- and right-hand sides are equal for every $f \in L_2(0,T;V^*)$ and $v \in H$. Hence $v \mapsto e^{-(T-\cdot)A^*}v$ is the adjoint of y_f.

Applying Corollary 2 to A^* for $t = 0$, a change of variables yields for every $v \in H$,

$$|v|^2 \leq \frac{C_5}{T}\int_0^T \|e^{-(T-s)A^*}v\|^2\,ds. \tag{96}$$

This estimate from below of the adjoint is equivalent to closedness of the range of y_f, as the range is dense by (94). This follows from the Closed Range Theorem; cf. ([26], Theorem 3.1) for a general result on this. $\square$

Remark 11. *The Banach spaces $L_2(0,T;V)$, $L_2(0,T;V^*)$ are in duality, and $L_2(0,T;V)^*$ identifies with $L_2(0,T;V^*)$: for each $\Lambda \in L_2(0,T;V)^*$ the inner product a_{Re} and Riesz' theorem yield $h \in L_2(0,T;V)$ that for $g \in L_2(0,T;V)$ fulfils $\langle \Lambda, g \rangle = \int_0^T a_{\mathrm{Re}}(h,g)\,dt$; so $\langle \Lambda, g \rangle = \int_0^T \langle f, g \rangle\,dt$ for $f = \frac{1}{2}(\mathcal{A}+\mathcal{A}')h$ in $L_2(0,T;V^*)$; cf. (23) and (25).*

The surjectivity of y_f can be shown in important cases using an explicit construction, which is of interest in control theory (cf. Remark 12), and given here for completeness:

Proposition 6. *If $A^* = A$ and A^{-1} is compact, every $v \in H$ equals y_f for some computable $f \in L_2(0,T;V^*)$.*

Proof. Fact 1 yields an ortonormal basis $(e_n)_{n\in\mathbb{N}}$ so that $Ae_n = \lambda_n e_n$, hence any v in H fulfils $v = \sum_j \alpha_j e_j$ with $\sum_j |\alpha_j|^2 < \infty$. By Fact 2 every $f \in L_2(0,T;V^*)$ has an expansion

$$f(t) = \sum_{j=1}^{\infty} \beta_j(t) e_j = \sum_{j=1}^{\infty} \langle f(t), e_j \rangle e_j \tag{97}$$

converging in V^* for t a.e. Since $e^{-(T-s)\mathcal{A}} e_j = e^{-(T-s)\lambda_j} e_j$, cf. Remark 2, such f fulfill

$$y_f = \int_0^T e^{-(T-s)\mathcal{A}} f(s)\,ds = \sum_{j=1}^{\infty} e^{-T\lambda_j} \Big(\int_0^T \beta_j(s) e^{s\lambda_j}\,ds \Big) e_j. \tag{98}$$

Hence $y_f = v$ is equivalent to the validity of $\int_0^T \beta_j(s) e^{s\lambda_j}\,ds = \alpha_j e^{T\lambda_j}$ for $j \in \mathbb{N}$. So if, in terms of some $\theta_j \in]0,1[$ to be determined, we take the coefficients of $f(t)$ as

$$\beta_j(t) = k_j \mathbf{1}_{[\theta_j T, T]}(t) \exp(t(\sqrt{\lambda_j} - \lambda_j)), \tag{99}$$

then the condition will be satisfied if and only if $k_j = \alpha_j e^{T\lambda_j} \sqrt{\lambda_j} (e^{T\sqrt{\lambda_j}} - e^{\theta_j T \sqrt{\lambda_j}})^{-1}$.

Moreover, using the equivalent norm $|||\cdot|||_*$ on V^* in Fact 2,

$$\|f\|^2_{L_2(0,T;V^*)} = \int_0^T |||f(t)|||_*^2\,dt = \sum_{j=1}^{\infty} \lambda_j^{-1} \int_0^T |\beta_j(t)|^2\,dt. \tag{100}$$

Therefore f is in $L_2(0,T;V^*)$ whenever $\int_0^T |\beta_j|^2\,dt \le C\lambda_j |\alpha_j|^2$ holds eventually for some $C > 0$, and here a direct calculation gives

$$\int_0^T \frac{|\beta_j|^2}{|k_j|^2}\,dt = \frac{e^{2T(\sqrt{\lambda_j}-\lambda_j)} - e^{2\theta_j T(\sqrt{\lambda_j}-\lambda_j)}}{2\sqrt{\lambda_j} - 2\lambda_j} = \frac{e^{2T\sqrt{\lambda_j}}(e^{2T(1-\theta_j)(\lambda_j-\sqrt{\lambda_j})} - 1)}{2e^{2T\lambda_j}(\lambda_j - \sqrt{\lambda_j})}. \tag{101}$$

So in view of the expression for k_j, the quadratic integrability of f follows if the θ_j can be chosen so that the above numerator is estimated by $C(\lambda_j - \sqrt{\lambda_j})(e^{T\sqrt{\lambda_j}} - e^{\theta_j T\sqrt{\lambda_j}})^2$ with C independent of $j \ge J$ for a suitable J, or more simply if

$$e^{2T(1-\theta_j)(\lambda_j-\sqrt{\lambda_j})} - 1 \le C(\lambda_j - \sqrt{\lambda_j})(1 - e^{-(1-\theta_j)T\sqrt{\lambda_j}})^2. \tag{102}$$

We may take J so that $\lambda_j > 3$ for all $j \geq J$, since at most finitely many eigenvalues do not fulfill this. Then $\theta_j := 1 - (\lambda_j - \sqrt{\lambda_j})^{-1}$ belongs to $]0,1[$, and the above is reduced to

$$\exp(2T) - 1 \leq C(\lambda_j - \sqrt{\lambda_j})(1 - \exp(-\frac{T}{\sqrt{\lambda_j} - 1}))^2. \tag{103}$$

Applying the Mean Value Theorem to exp on $[-\frac{T}{\sqrt{\lambda_j}-1}, 0]$, we obtain the inequality

$$(\lambda_j - \sqrt{\lambda_j})(1 - \exp(\frac{-T}{\sqrt{\lambda_j} - 1}))^2 \geq \exp(-\frac{2T}{\sqrt{3} - 1})\frac{T^2\lambda_j}{\lambda_j - \sqrt{\lambda_j}} > \exp(-4T)T^2 > 0. \tag{104}$$

Hence (103) is fulfilled for $C = \exp(6T)/T^2$. □

Remark 12. *In the above proof* $\operatorname{supp}\beta_j \subset [\theta_j T, T]$*, so the given* v *can be attained by* y_f *by arranging the coefficients* β_j *in each dimension* successively *as time approaches* T*, as* $\theta_j \nearrow 1$ *follows in* (99) *by counting the eigenvalues so that* $\lambda_j \nearrow \infty$*. This can even be postponed to any given* $T_0 < T$*, for* $\operatorname{supp}\beta_j \subset [T_0, T]$ *holds whenever* $\theta_j T \geq T_0$*, and we may reset to* $\theta_j = T_0/T$ *and adjust the* k_j *accordingly, for the finitely many remaining* j*. Both themes may be of interest in infinite dimensional control theory.*

In order to isolate $u(0)$ in (91), it will of course be decisive that the operator e^{-TA} has an inverse, as was shown for general analytic semigroups in Proposition 1.

For our Lax–Milgram operator A with analytic semigroup e^{-tA} generated by $\mathbf{A} = -A$, it is the symbol e^{tA} that denotes the inverse, consistent with the sign convention in (41). Hence the properties of e^{tA} can be read off from Proposition 2, where (43) gives

$$e^{-tA}e^{TA} \subset e^{(T-t)A} \qquad \text{for } 0 \leq t \leq T. \tag{105}$$

Moreover, it is decisive for the interpretation of the compatibility conditions in Section 4.2 below to know that the domain inclusions in Proposition 2 are strict. We include a mild sufficient condition along with a characterisation of the domain $D(e^{tA})$.

Proposition 7. *If* H *has an orthonormal basis of eigenvectors* $(e_j)_{j\in\mathbb{N}}$ *of* A *so that the corresponding eigenvalues fulfil* $\operatorname{Re}\lambda_j \to \infty$ *for* $j \to \infty$*, then the inclusions in* (44) *are both* strict*, and* $D(e^{tA})$ *is the completion of* $\operatorname{span}(e_j)_{j\in\mathbb{N}}$ *with respect to the graph norm,*

$$\|x\|^2_{D(e^{tA})} = \sum_{j=1}^{\infty}(1 + e^{2\operatorname{Re}\lambda_j t})|(x \mid e_j)|^2. \tag{106}$$

The domain $D(e^{tA})$ *equals the subspace* $S \subset H$ *in which the right-hand side is finite.*

Proof. If $x \in S$ the vector $v = \sum_{j=1}^{\infty} e^{\lambda_j t}(x \mid e_j)e_j$ is well defined in H, and with methods from Remark 2 it follows that $e^{-tA}v = x$; i.e., $x \in D(e^{tA})$.

Conversely, for $x \in D(e^{tA})$ there is a vector $y \in H$ such that $x = e^{-tA}y = \sum_{j=1}^{\infty}(y \mid e_j)e^{-t\lambda_j}e_j$. That is, $e^{\lambda_j t}(x \mid e_j) = (y \mid e_j) \in \ell_2$, so $x \in S$. Then $|e^{tA}x|^2 = \sum e^{2\operatorname{Re}\lambda_j t}|(x \mid e_j)|^2$ yields (106).

Now any $x \in D(e^{t'A})$ is also in $D(e^{tA})$ for $t < t'$, since $\operatorname{Re}\lambda_j > 0$ holds in (106) for all j by V-ellipticity. As $\operatorname{Re}\lambda_j \to \infty$, we may choose a subsequence so that $\operatorname{Re}\lambda_{j_n} > n$ and set

$$x = \sum_{n=1}^{\infty} \frac{1}{n} e^{-\lambda_{j_n} t} e_{j_n}. \tag{107}$$

Here $x \in D(e^{tA})$ as it is in S by construction for $t \geq 0$; but not in $D(e^{t'A})$ for $t' > t$ as

$$\sum_{j=1}^{\infty} e^{2\operatorname{Re}\lambda_j t'} |(x \,|\, e_j)|^2 = \sum_{n=1}^{\infty} e^{2\operatorname{Re}\lambda_{j_n}(t'-t)} \frac{1}{n^2} > \sum_{n=1}^{\infty} \frac{e^{2n(t'-t)}}{n^2} = \infty. \tag{108}$$

Furthermore, using orthogonality, it follows for any $x \in D(e^{tA})$ that, for $N \to \infty$,

$$\|x - \sum_{j \leq N} (x \,|\, e_j) e_j\|^2_{D(e^{tA})} = \sum_{J > N} (1 + e^{2\operatorname{Re}\lambda_j t}) |(x \,|\, e_j)|^2 \to 0. \tag{109}$$

Hence the space $D(e^{tA})$ has $\operatorname{span}(e_j)_{j \in \mathbb{N}}$ as a dense subspace. That is, the completion of the latter with respect to the graph norm identifies with the former. □

After this study of the map y_f, the injectivity of the operator e^{-tA} and the domain $D(e^{tA})$, cf. Propositions 1, 2, 5 and 7, we address the final value problem (90) by solving (91) for the vector $u(0)$. This is done by considering the map

$$u(0) \mapsto e^{-TA} u(0) + y_f. \tag{110}$$

This is composed of the bijection e^{-TA} and a translation by the vector y_f, hence is bijective from H to the affine space $R(e^{-TA}) + y_f$. In fact, using (41), inversion gives

$$u(0) = e^{TA} \big(u(T) - \int_0^T e^{-(T-s)\mathcal{A}} f(s)\, ds\big) = e^{TA} (u(T) - y_f). \tag{111}$$

This may be summed up thus:

Theorem 6. *For the set of solutions u in X of the differential equation $(\partial_t + \mathcal{A})u = f$ with fixed data $f \in L_2(0,T;V^*)$, the Formulas (91) and (111) give a bijective correspondence between the initial states $u(0)$ in H and the terminal states $u(T)$ in $y_f + D(e^{TA})$.*

In view of the linearity, the affine space $y_f + D(e^{TA})$ might seem surprising. However, a suitable reinterpretation gives the compatibility condition introduced in the next section.

4.2. Well-Posedness of the Final Value Problem

Since $R(e^{TA}) \subset H$, the initial state in (111) can be inserted into Formula (76), so any solution u of (90) must satisfy

$$u(t) = e^{-tA} e^{TA} (u_T - y_f) + \int_0^t e^{-(t-s)\mathcal{A}} f(s)\, ds. \tag{112}$$

Here one could contract the first term a bit, as $e^{-tA}e^{TA} \subset e^{(T-t)A}$ by (105). But we refrain from this because $e^{-tA}e^{TA}$ rather obviously applies to $u_T - y_f$ if and only if this vector belongs to $D(e^{TA})$ — and the following theorem corroborates that this is equivalent to the unique solvability in X of the final value problem (90):

Theorem 7. *Let V be a separable Hilbert space contained algebraically, topologically and densely in H, and let A be the Lax–Milgram operator defined in H from a bounded V-elliptic sesquilinear form a, and having bounded extension $\mathcal{A}\colon V \to V^*$. For given $f \in L_2(0,T;V^*)$ and $u_T \in H$, the condition*

$$u_T - y_f \in D(e^{TA}) \tag{113}$$

is necessary and sufficient for the existence of some $u \in X$, cf. (60), that solves the final value problem (90). Such a function u is uniquely determined and given by (112), where all terms belong to X as functions of t.

Proof. When (90) has a solution $u \in X$, then u_T is reachable from the initial state $u(0)$ determined from the bijection in Theorem 6, which gives that $u_T - y_f = e^{-TA}u(0) \in D(e^{TA})$. Hence (113) is necessary and (112) follows by insertion, as explained prior to (112). Uniqueness is obvious from the right-hand side of (112).

When u_T, f fulfill (113), then $u_0 = e^{TA}(u_T - y_f)$ defines a vector in H, so Theorem 4 yields a function $u \in X$ solving $(\partial_t + \mathcal{A})u = f$ and $u(0) = u_0$. According to Theorem 6 this u has final state $u(T) = e^{-TA}e^{TA}(u_T - y_f) + y_f = u_T$, hence solves (90).

Finally, the fact that the integral in (112) defines a function in X follows at once from Theorem 5, for it states that it equals the solution in X of $\tilde{u}' + \mathcal{A}\tilde{u} = f$, $\tilde{u}(0) = 0$. Since $u \in X$ in (112), also $e^{-tA}e^{TA}(u_T - y_f)$ is a function in X. $\square$

Remark 13. *When (f, u_T) fulfils (113), then (111) yields that $u_T - y_f = e^{-TA}u(0)$.*

Remark 14. *Writing condition (113) as $u_T = e^{-TA}u(0) + y_f$, cf. Remark 13, this part of Theorem 7 is natural inasmuch as each set of admissible terminal data u_T are in effect a sum of the terminal state, $e^{-TA}u(0)$, of the semi-homogeneous initial value problem (53) with $f = 0$ and of the terminal state y_f of the semi-homogeneous problem (53) with $u(0) = 0$. Moreover, the u_T fill at least a dense set in H, as for fixed $u(0)$ this follows from Proposition 5; for fixed f from the density of $D(e^{TA})$ seen prior to Proposition 2.*

Remark 15. *To elucidate the criterion $u_T - y_f \in D(e^{TA})$ in formula (113) of Theorem 7, we consider the matrix operator $P_{\mathcal{A}} = \begin{pmatrix} \partial_t + \mathcal{A} \\ r_T \end{pmatrix}$, with r_T denoting restriction at $t = T$, and the "forward" map $\Phi(f, u_T) = u_T - y_f$, which by (61) and Proposition 5 give bounded operators*

$$X \xrightarrow{P_{\mathcal{A}}} \begin{matrix} L_2(0,T;V^*) \\ \oplus \\ H \end{matrix} \xrightarrow{\Phi} H. \tag{114}$$

Then, in terms of the range $R(P_{\mathcal{A}})$, clearly (90) has a solution if and only if $\begin{pmatrix} f \\ u_T \end{pmatrix} \in R(P_{\mathcal{A}})$, so the compatibility condition (113) means that $R(P_{\mathcal{A}}) = \Phi^{-1}(D(e^{TA})) = D(e^{TA}\Phi)$.

The paraphrase at the end of Remark 15 is convenient for the choice of a useful norm on the data. Indeed, we now introduce the space of admissible data $Y = D(e^{TA}\Phi)$, i.e.,

$$Y = \Big\{ (f, u_T) \in L_2(0,T;V^*) \oplus H \;\Big|\; u_T - y_f \in D(e^{TA}) \Big\}, \tag{115}$$

endowed with the graph norm on $D(e^{TA}\Phi)$ given by

$$\|(f, u_T)\|_Y^2 = |u_T|^2 + \|f\|^2_{L_2(0,T;V^*)} + |e^{TA}(u_T - y_f)|^2. \tag{116}$$

Using the equivalent norm $\|\!|\cdot|\!\|_*$ from (26) for V^*, the above is induced by the inner product

$$(u_T \,|\, v_T) + \int_0^T (f(s) \,|\, g(s))_{V^*}\, ds + (e^{TA}(u_T - y_f) \,|\, e^{TA}(v_T - y_g)). \tag{117}$$

This space Y is *complete*: as Φ in Remark 15 is bounded, the composite map $e^{TA}\Phi$ is a closed operator from $L_2(0,T;V^*) \oplus H$ to H, so its domain $D(e^{TA}\Phi) = Y$ is complete with respect to the graph norm given in (116). Hence Y is a Hilbert(-able) space—but we shall often just work with the equivalent norm on the Banach space Y obtained by using simply $\|\cdot\|_*$ on V^*.

Moreover, the norm in (116) also leads to continuity of the solution operator for (90):

Theorem 8. *The solution $u \in X$ in Theorem 7 depends continuously on the data (f, u_T) in the Hilbert space Y in (115), or equivalently, for some constant c we have*

$$\int_0^T \|u(t)\|^2\,dt + \sup_{t\in[0,T]} |u(t)|^2 + \int_0^T \|\partial_t u(t)\|_*^2\,dt \\ \le |u_T|^2 + c\Big(\int_0^T \|f(t)\|_*^2\,dt + \big|e^{TA}(u_T - \int_0^T e^{-(T-t)\mathcal{A}} f(t)\,dt)\big|^2\Big). \tag{118}$$

Another equivalent norm on the Hilbert space Y is obtained by omitting the term $|u_T|^2$.

Proof. This follows from Corollary 1 by inserting $u_0 = e^{TA}(u_T - y_f)$ from (111) into (64), for this gives $\|u\|_X^2 \le c|e^{TA}(u_T - y_f)|^2 + c\|f\|^2_{L_2(0,T;V^*)}$, where one can add $|u_T|^2$. Conversely the boundedness of y_f and e^{-TA} yield that $|u_T|^2 \le c\|f\|^2 + c|e^{TA}(u_T - y_f)|^2$. □

Of course, Theorems 7 and 8 together mean that the final value problem in (90) is well posed in the spaces X and Y.

5. The Heat Equation With Final Data

To apply the theory in Section 4, we treat the heat equation and its final value problem. In the sequel Ω stands for a smooth, open bounded set in $\mathbb{R}^n$, $n \ge 2$ as described in ([20], Appendix C). In particular Ω is locally on one side of its boundary $\Gamma := \partial\Omega$.

For such sets we consider the problem of finding the u satisfying

$$\begin{cases} \partial_t u(t,x) - \Delta u(t,x) = f(t,x) & \text{in } Q :=]0,T[\times\Omega, \\ \gamma_0 u(t,x) = g(t,x) & \text{on } S :=]0,T[\times\partial\Omega, \\ r_T u(x) = u_T(x) & \text{at } \{T\}\times\Omega. \end{cases} \tag{119}$$

Hereby the trace of functions on Γ is written in the operator notation $\gamma_0 u = u|_\Gamma$; similarly we also use γ_0 for traces on S. r_T denotes the trace operator at $t = T$.

We shall also use $H_0^1(\Omega)$, which is the subspace obtained by closing $C_0^\infty(\Omega)$ in the Sobolev space $H^1(\Omega)$. Dual to this one has $H^{-1}(\Omega)$, which identifies with the set of restrictions to Ω from $H^{-1}(\mathbb{R}^n)$, endowed with the infimum norm. The reader is referred to Chapter 6 and Remark 9.4 in [20] for the spaces $H^s(\mathbb{R}^n)$ and the infimum norm.

5.1. The Boundary Homogeneous Case

In case $g \equiv 0$ in (119), the consequences of the abstract results in Section 4.2 are straightforward to account for. Indeed, with

$$V = H_0^1(\Omega), \quad H = L_2(\Omega), \quad V^* = H^{-1}(\Omega), \tag{120}$$

the boundary condition $\gamma_0 u = 0$ is imposed via the condition that $u(t) \in V$ for all t, or rather through use of the Dirichlet realization of the Laplacian $-\Delta_{\gamma_0}$ (denoted by $-\Delta_D$ in the introduction), which is the Lax–Milgram operator A induced by the triple $(L_2(\Omega), H_0^1(\Omega), s)$ for

$$s(u,v) = \sum_{j=1}^n (\partial_j u \,|\, \partial_j v)_{L_2(\Omega)}. \tag{121}$$

In fact, the Poincaré inequality yields that the form $s(u,v)$ is $H_0^1(\Omega)$-elliptic, and as it is symmetric too, $A = -\Delta_{\gamma_0}$ is a selfadjoint unbounded operator in $L_2(\Omega)$, with $D(-\Delta_{\gamma_0}) \subset H_0^1(\Omega)$.

Hence the operator $-A = \Delta_{\gamma_0}$ generates an analytic semigroup $e^{t\Delta_{\gamma_0}}$ in $\mathbb{B}(L_2(\Omega))$; the bounded extension $-\mathcal{A} = \Delta\colon H_0^1(\Omega) \to H^{-1}(\Omega)$ generates the analytic semigroup $e^{-t\mathcal{A}} = e^{t\Delta}$ on $H^{-1}(\Omega)$; cf. Lemma 4. Consistently with Section 4.1 we also set $(e^{t\Delta_{\gamma_0}})^{-1} = e^{-t\Delta_{\gamma_0}}$.

For the homogeneous problem with $g = 0$ in (119) we have the solution and data spaces

$$X_0 = L_2(0,T;H_0^1(\Omega)) \bigcap C([0,T];L_2(\Omega)) \bigcap H^1(0,T;H^{-1}(\Omega)), \tag{122}$$

$$Y_0 = \left\{ (f,u_T) \in L_2(0,T;H^{-1}(\Omega)) \oplus L_2(\Omega) \;\middle|\; u_T - y_f \in D(e^{-T\Delta_{\gamma_0}}) \right\}. \tag{123}$$

Here, with y_f as the usual integral (cf. (125) below), the data norm in (116) amounts to

$$\|(f,u_T)\|_{Y_0}^2 = \int_0^T \|f(t)\|_{H^{-1}(\Omega)}^2\,dt + \int_\Omega (|u_T|^2 + |e^{-T\Delta_{\gamma_0}}(u_T - y_f)|^2)\,dx. \tag{124}$$

From Theorems 7 and 8 we may now read off the following result, which is a novelty even though the problem is classical:

Theorem 9. *Let $A = -\Delta_{\gamma_0}$ be the Dirichlet realization of the Laplacian in Ω and $\mathcal{A} = -\Delta$ its extension, as introduced above. When $g = 0$ in the final value problem (119) and $f \in L_2(0,T;H^{-1}(\Omega))$, $u_T \in L_2(\Omega)$, then there exists a solution u in X_0 of (119) if and only if the data (f,u_T) are given in Y_0, i.e., if and only if*

$$u_T - \int_0^T e^{-(T-s)\mathcal{A}} f(s)\,ds \quad \textit{belongs to} \quad D(e^{-T\Delta_{\gamma_0}}). \tag{125}$$

In the affirmative case, such u are uniquely determined in X_0 and fulfil the estimate $\|u\|_{X_0} \le c\|(f,u_T)\|_{Y_0}$. Furthermore the difference in (125) equals $e^{T\Delta_{\gamma_0}}u(0)$ in $L_2(\Omega)$.

Remark 16. *For $A = -\Delta_{\gamma_0}$ one has the equivalent norms in Facts 1, 2 and the characterisation of $D(e^{-T\Delta_{\gamma_0}})$ in Proposition 7. This is a classical consequence of the compact embedding of $H_0^1(\Omega)$ into $L_2(\Omega)$ for bounded sets Ω (e.g., ([20], Theorem 8.2)). Thus one obtains for $f = 0$, $g = 0$ the situation described in the introduction, where the space of final data, normed by $|||u_T|||$, via Proposition 7 is seen to be $D(e^{-T\Delta_{\gamma_0}})$ with equivalent norms. As the completed solution space $\overline{\mathcal{E}}$ in the introduction one may take the Banach space $\overline{\mathcal{E}} = X_0$, cf. Theorem 9.*

5.2. The Inhomogeneous Case

For non-zero data, i.e., when $g \neq 0$ on S, cf. (119), one may of course try to reduce to an equivalent homogeneous problem by choosing a function w so that $\gamma_0 w = g$ on the surface S. Here we recall the classical

Lemma 6. *$\gamma_0\colon H^1(Q) \to H^{1/2}(S)$ is a continuous surjection having a bounded right inverse $\tilde{K}_0$, so $w = \tilde{K}_0 g$ maps every $g \in H^{1/2}(S)$ to $w \in H^1(Q)$ fulfilling $\gamma_0 w = g$ and*

$$\|w\|_{H^1(Q)} \le c\|g\|_{H^{1/2}(S)}. \tag{126}$$

Lacking a reference with details, we note that the lemma is well known for sets like Ω, hence for smooth open bounded sets $\Omega_1 \subset \mathbb{R}^{n+1}$ with operators γ_{0,Ω_1} and $\tilde{K}_{0,\Omega_1}$; cf. Theorem B.1.9 in [22] or Theorem 9.5 in [20] for the flat case. In particular, one can stretch Q to $]-2T,2T[\times\Omega$ and attach rounded ends in a smooth way to obtain a set $\Omega_1 \subset]-3T,3T[\times\Omega$ equal to Q for $0 < t < T$. Here $H^1(Q) = r_Q H^1(\Omega_1)$ is a classical result, when the latter space of restrictions to Q has the infimum norm. While $H^s(\partial\Omega_1)$ is defined using local coordinates in a standard way, cf. ([20], Formula (8.10)), the Sobolev space $H^s(S)$ on the surface S can be defined as the set of restrictions $r_S H^s(\partial\Omega_1)$. When $r_S\tilde{g} = g$, then $\tilde{K}_0 g = r_Q \tilde{K}_{0,\Omega_1}\tilde{g}$ defines the desired operator $\tilde{K}_0$, as γ_{0,Ω_1} acts as γ_0 in Q.

Remark 17. *The norm in $H^s(S)$ can be chosen so that this is a Hilbert space; cf. ([20], Formula (8.10)). However, Sobolev spaces on smooth surfaces is a vast subject, requiring so-called distribution densities as explained in ([22], Section 6.3). We refer the reader to ([20], Section 8.2) for a short introduction to this subject; as there, we prefer a more intuitive approach (exploiting the surface measure on Ω_1) but skip details. A systematic exposition of this framework can be found in ([27], Section 4), albeit in a general L_p-setting with mixed-norms leading to anisotropic Triebel–Lizorkin spaces $F^{s,\vec{a}}_{\vec{p},q}(S)$ on the curved boundary, which in general are the correct boundary data spaces for parabolic problems with different integrability properties in space and time, as noted in [28]; cf. the discussion of the heat equation in ([27], Section 6.5) and the more detailed account in ([29], Chapter 7).*

However, when splitting the solution of (119) as $u = v + w$ for w as in Lemma 6, then v should satisfy (119) with data $(\tilde{f}, 0, \tilde{u}_T)$,

$$\tilde{f} = f - (\partial_t w - \Delta w), \qquad \tilde{u}_T = u_T - r_T w. \tag{127}$$

At first glance one might therefore think that w is inconsequential for the compatibility condition (125), for $\tilde{u}_T - y_{\tilde{f}}$ there equals the usual term $u_T - y_f$ minus $r_T w - y_{\partial_t w - \Delta w}$, where the latter seemingly belongs to $D(e^{-T\Delta_{\gamma_0}})$ as the pair $(\partial_t w - \Delta w, r_T w)$ could seem to be a vector in the range of the operator $P_{-\Delta}$ in Remark 15.

But obviously this is not the case, because the function w is outside the domain X_0 of $P_{-\Delta}$. Indeed, $w \in L_2(0,T; H^1(\Omega))$ and has $\gamma_0 w = g \not\equiv 0$ in the non-homogeneous case, whence $w \notin L_2(0,T; H^1_0(\Omega))$. So one might think it would be necessary to discuss homogeneous problems with larger solution spaces $\tilde{X}_0$ than X_0.

We propose to circumvent these difficulties by applying Lemma 6 to the corresponding linear *initial* value problem instead, since in the present spaces of low regularity there is no compatibility condition needed for this:

$$\begin{cases} \partial_t u - \Delta u = f & \text{in } Q, \\ \gamma_0 u = g & \text{on } S, \\ r_0 u = u_0 & \text{at } \{0\} \times \Omega. \end{cases} \tag{128}$$

More precisely, we shall analogously to Section 4 obtain a bijection $u(0) \leftrightarrow u(T)$ between initial and final states by establishing a solution formula as in Theorem 5. (For general background material on (128) the reader could consult Section III.6 in [30], and for the fine theory including compatibility conditions we refer to [15].)

Analogously to Theorem 4 and Corollary 1, we depart from well-posedness of (128). This is well known *per se*, but we need to briefly review the explanation in order to account later for the decisive existence of an improper integral showing up when $g \neq 0$ in (119).

Since the solutions now take values in the full space $H^1(\Omega)$, we shall in this section denote the solution space by X_1. It is given by

$$X_1 = L_2(0,T; H^1(\Omega)) \bigcap C([0,T]; L_2(\Omega)) \bigcap H^1(0,T; H^{-1}(\Omega)), \tag{129}$$

and X_1 is a Banach space when normed analogously to (61),

$$\|u\|_{X_1} = (\|u\|^2_{L_2(0,T;H^1(\Omega))} + \sup_{0 \le t \le T} \|u(t)\|^2_{L_2(\Omega)} + \|u\|^2_{H^1(0,T;H^{-1}(\Omega))})^{1/2}. \tag{130}$$

As H^1, H^{-1} are not dual on Ω, the redundancy in Remark 5 does not extend to the term $\sup_{[0,T]} \|u\|_{L_2}$ above.

Proposition 8. *The heat initial value problem* (128) *has a unique solution $u \in X_1$ for given data $f \in L_2(0,T;H^{-1}(\Omega))$, $g \in H^{1/2}(S)$, $u_0 \in L_2(\Omega)$, and there is an estimate*

$$\|u\|^2_{X_1} \le c(\|u_0\|^2_{L_2(\Omega)} + \|f\|^2_{L_2(0,T;H^{-1}(\Omega))} + \|g\|^2_{H^{1/2}(S)}). \tag{131}$$

Proof. With $w = \tilde{K}_0 g$ as in Lemma 6, we write $u = v + w$ for some $v \in X_1$ solving (128) for data

$$\tilde{f} = f - (\partial_t - \Delta)w, \qquad \tilde{g} = 0, \qquad \tilde{u}_0 = u_0 - w(0). \tag{132}$$

Here $w(0)$ is well defined, as $w \in H^1(Q)$ implies $w \in C([0,T];L_2(\Omega))$, by an application of Lemma 1. That w even is in X_1 results from the easy estimates, where $I =]0,T[$,

$$\|w'\|^2_{L_2(I;H^{-1})} + \|\Delta w\|^2_{L_2(I;H^{-1})} \le \|w\|^2_{H^1(I;L_2)} + c\|w\|^2_{L_2(I;H^1)} \le c\|w\|^2_{H^1(Q)}. \tag{133}$$

This moreover yields that $\tilde{f} \in L_2(0,T;H^{-1}(\Omega))$, and $\tilde{u}_0 \in L_2(\Omega)$, so by Theorem 4, the boundary homogeneous problem for v has a solution in X_0; cf. (122). Hence (128) has the solution $u = v + w$ in X_1; and by linearity this is unique in view of Theorem 4.

Inspecting the above arguments, we first note that by (57),

$$\sup_{0\le t\le T} \|w(t)\|_{L_2(\Omega)} \le c(\|w\|_{L_2(0,T;L_2(\Omega))} + \|\partial_t w\|_{L_2(0,T;L_2(\Omega))}) \le c\|w\|^2_{H^1(Q)}, \tag{134}$$

so the estimate (133) can be sharpened to $\|w\|^2_{X_1} \le c\|w\|^2_{H^1(Q)}$. Now Corollary 1 gives

$$\begin{aligned}\|u\|^2_{X_1} &\le 2(\|v\|^2_{X_0} + \|w\|^2_{X_1}) \le c(\|\tilde{u}_0\|^2_{L_2(\Omega)} + \|\tilde{f}\|^2_{L_2(0,T;H^{-1}(\Omega))} + \|w\|^2_{X_1})\\ &\le c(\|u_0\|^2_{L_2(\Omega)} + \|f\|^2_{L_2(0,T;H^{-1}(\Omega))} + \|(\partial_t - \Delta)w\|^2_{L_2(0,T;H^{-1})} + \|w\|^2_{H^1(Q)})\end{aligned} \tag{135}$$

which via (133) and (126) entails the stated estimate (131). □

As a crucial addendum, we may apply Theorem 5 directly to the function v constructed during the above proof and then substitute $v = u - w$ to derive that

$$u(t) = w(t) + e^{t\Delta_{\gamma_0}}(u_0 - w(0)) + \int_0^t e^{-(t-s)\mathcal{A}}(f - (\partial_s - \Delta)w)\,ds. \tag{136}$$

This formula for the u solving the inhomogeneous final value problem applies especially for $t = T$, but we shall keep t in $[0,T]$ to deduce a formula for its solution.

Our strategy in the following will be to simplify the contributions from w, and ultimately to reintroduce the boundary data g instead of w. To do so, we apply the Leibniz rule in Proposition 3 to our function w in $H^1(0,t;L_2(\Omega))$ and get

$$\partial_s(e^{(t-s)\Delta_{\gamma_0}} w(s)) = e^{(t-s)\Delta_{\gamma_0}} \partial_s w(s) - \Delta_{\gamma_0} e^{(t-s)\Delta_{\gamma_0}} w(s). \tag{137}$$

As the first inconvenience, Δ_{γ_0} does not commute with the semigroup, since w as an element of $H^1 \setminus H^1_0$ belongs to neither the domain of the realization $-\Delta_{\gamma_0}$, nor to that of $\mathcal{A}$.

Secondly, the right-hand side is only integrable on $[0, t-\varepsilon]$ for $\varepsilon > 0$, as the last term has a singularity at $s = t$; cf. Theorem 3. As a remedy, we may use the improper Bochner integral

$$⨍_0^t \Delta_{\gamma_0} e^{(t-s)\Delta_{\gamma_0}} w(s)\,ds = \lim_{\varepsilon\to 0} \int_0^{t-\varepsilon} \Delta_{\gamma_0} e^{(t-s)\Delta_{\gamma_0}} w(s)\,ds. \tag{138}$$

Lemma 7. *For every $w \in H^1(Q)$ the limit (138) exists in $L_2(\Omega)$ and*

$$w(t) - e^{t\Delta_{\gamma_0}} w(0) = \int_0^t e^{(t-s)\Delta_{\gamma_0}} \partial_s w(s)\, ds - \fint_0^t \Delta_{\gamma_0} e^{(t-s)\Delta_{\gamma_0}} w(s))\, ds. \tag{139}$$

Proof. As $e^{t\Delta_{\gamma_0}}$ is uniformly bounded according to Theorem 3 and $w \in C([0,T], L_2(\Omega))$ was seen in the above proof, bilinearity gives that in $L_2(\Omega)$,

$$e^{(t-(t-\varepsilon))\Delta_{\gamma_0}} w(t-\varepsilon) \to w(t) \quad \text{for } \varepsilon \to 0. \tag{140}$$

Moreover, integration of both sides in (137) gives, cf. Lemma 1,

$$[e^{(t-s)\Delta_{\gamma_0}} w(s)]_{s=0}^{s=t-\varepsilon} = \int_0^{t-\varepsilon} (-\Delta_{\gamma_0}) e^{(t-s)\Delta_{\gamma_0}} w(s)\, ds + \int_0^{t-\varepsilon} e^{(t-s)\Delta_{\gamma_0}} \partial_s w(s)\, ds. \tag{141}$$

The left-hand side converges by (140), and by dominated convergence the rightmost term does so for $\varepsilon \to 0^+$ (through an arbitrary sequence), so also $\int_0^{t-\varepsilon} \Delta_{\gamma_0} e^{(t-s)\Delta_{\gamma_0}} w(s)\, ds$ converges in $L_2(\Omega)$ as claimed. Then (139) is the resulting identity among the limits. □

Identity (139) from the lemma applies directly in the solution formula (136), and because terms with $\partial_s w$ cancel, one obtains

$$u(t) = e^{t\Delta_{\gamma_0}} u_0 + \int_0^t e^{-(t-s)\mathcal{A}} f\, ds + \int_0^t e^{-(t-s)\mathcal{A}} \Delta w\, ds - \fint_0^t \Delta_{\gamma_0} e^{(t-s)\Delta_{\gamma_0}} w\, ds. \tag{142}$$

We shall reduce the difference of the last two integrals in order to reintroduce the boundary data g instead of w.

First we use that $\Delta = \mathcal{A}\mathcal{A}^{-1}\Delta$ on $H^1(\Omega)$ and write both terms as improper integrals,

$$-\fint_0^t \mathcal{A} e^{-(t-s)\mathcal{A}} (I - \mathcal{A}^{-1}\Delta) w(s)\, ds. \tag{143}$$

Here $Q = I - \mathcal{A}^{-1}\Delta$ is a well-known projection from the fine elliptic theory of the problem

$$-\Delta u = f, \quad \gamma_0 u = g. \tag{144}$$

Indeed, if this is treated via the matrix operator $\left(\begin{smallmatrix} -\Delta \\ \gamma_0 \end{smallmatrix}\right)$, which has an inverse in row form $\left(\begin{smallmatrix} -\mathcal{A}^{-1} & K_0 \end{smallmatrix}\right)$ that applies to the data $\left(\begin{smallmatrix} f \\ g \end{smallmatrix}\right)$, the basic composites appear in the two operator identities on $H^1(\Omega)$ and $H^{-1}(\Omega) \oplus H^{1/2}(\Gamma)$ respectively,

$$I = \begin{pmatrix} -\mathcal{A}^{-1} & K_0 \end{pmatrix} \begin{pmatrix} -\Delta \\ \gamma_0 \end{pmatrix} = \mathcal{A}^{-1}\Delta + K_0\gamma_0, \tag{145}$$

$$\begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} = \begin{pmatrix} -\Delta \\ \gamma_0 \end{pmatrix} \begin{pmatrix} -\mathcal{A}^{-1} & K_0 \end{pmatrix} = \begin{pmatrix} \Delta\mathcal{A}^{-1} & \Delta K_0 \\ -\gamma_0\mathcal{A}^{-1} & \gamma_0 K_0 \end{pmatrix}. \tag{146}$$

Thus we get from the first formula that $Q = I - \mathcal{A}^{-1}\Delta = K_0\gamma_0$ on $H^1(\Omega)$.

However, before we implement this, we emphasize that the simplicity of the Formulas (145) and (146) relies on a specific choice of K_0 explained in the following:

As $\mathcal{A} = \Delta|_{H_0^1}$ holds in the distribution sense, $P := \mathcal{A}^{-1}\Delta$ clearly fulfils $P^2 = P$, is bounded $H^1 \to H_0^1$ and equals I on H_0^1, so P is the projection onto $H_0^1(\Omega)$ along its null space, which evidently is the closed subspace of harmonic H^1-functions, namely

$$Z(-\Delta) = \{\, u \in H^1(\Omega) \mid -\Delta u = 0 \,\}. \tag{147}$$

Therefore H^1 is a direct sum,

$$H^1(\Omega) = H_0^1(\Omega) \dotplus Z(-\Delta). \tag{148}$$

We also let $Q = I - P$ denote the projection on $Z(-\Delta)$ along $H_0^1(\Omega)$, as from the context it can be distinguished from the time cylinder (also denoted by Q).

Since $\gamma_0 \colon H^1(\Omega) \to H^{1/2}(\Gamma)$ is surjective with H_0^1 as the null-space, it has an inverse K_0 on the complement $Z(-\Delta)$, which by the open mapping principle is bounded

$$K_0 \colon H^{1/2}(\Gamma) \to Z(-\Delta). \tag{149}$$

Hence $K_0 \colon H^{1/2}(\Gamma) \to H^1(\Omega)$ is a bounded right-inverse, i.e., $\gamma_0 K_0 = I_{H^{1/2}(\Gamma)}$. The rest of (146) follows at once. Moreover, since $\gamma_0 P = 0$,

$$K_0\gamma_0 = K_0\gamma_0(P + Q) = K_0\gamma_0 Q = I_{Z(-\Delta)}Q = Q, \tag{150}$$

which by definition of Q and P gives (145). (K_0 is known as a Poisson operator; these are amply discussed within the pseudo-differential boundary operator calculus in [31].)

Using this set-up we obtain:

Proposition 9. *If u denotes the unique solution to the initial boundary value problem (128) provided by Proposition 8, then u fulfils the identity*

$$u(t) = e^{t\Delta_{\gamma_0}}u_0 + \int_0^t e^{-(t-s)\mathcal{A}} f(s)\,ds - \fint_0^t \mathcal{A}e^{(t-s)\Delta_{\gamma_0}}K_0 g(s)\,ds, \tag{151}$$

where the improper integral converges in $L_2(\Omega)$ for every $t \in [0, T]$.

Proof. Because of (150) we may write $(I - \mathcal{A}^{-1}\Delta)w = Qw = K_0\gamma_0 w = K_0 g$ when $\gamma_0 w = g$, and when this is applied in (143), the solution formula (142) simplifies to (151). $\square$

For $t = T$ the second term in (151) gives back $y_f = \int_0^T e^{-(T-s)\mathcal{A}} f(s)\,ds$ from Section 4. However, the *full* influence on $u(T)$ from the boundary data g is collected in the third term as

$$z_g = \fint_0^T \mathcal{A}e^{(T-s)\Delta_{\gamma_0}}K_0 g(s)\,ds. \tag{152}$$

That the map $g \mapsto z_g$ is well defined is clear by taking $t = T$ in Proposition 9; this is a non-trivial result. The map is linear by the calculus of limits. In case $f = 0$, $u_0 = 0$ it is seen from (151) that $z_g = u(T)$, so obviously $\|z_g\|_{L_2(\Omega)} \le \sup_t \|u(t)\|_{L_2(\Omega)}$, which in turn is estimated by $c\|g\|_{H^{1/2}(S)}$ using Proposition 8. This proves

Lemma 8. *The linear operator $g \mapsto z_g$ is bounded $H^{1/2}(S) \to L_2(\Omega)$.*

Finally, from Proposition 9, we conclude for an arbitrary solution in X_1 of the heat equation $u' - \Delta u = f$ with $\gamma_0 u = g$ on S that

$$u(T) = e^{T\Delta_{\gamma_0}}u(0) + y_f - z_g. \tag{153}$$

Therefore we also here have a bijection $u(0) \leftrightarrow u(T)$, for the above breaks down to application of the bijection $e^{T\Delta_{\gamma_0}}$, cf. Proposition 1, and a translation in $L_2(\Omega)$ by the fixed vector $y_f - z_g$.

We are now ready to obtain the unique solvability of the inhomogeneous final value problem (119). Our result for this is similar to the abstract Theorem 7 (as is its proof), except for the important clarification that the boundary data g *do* appear in the compatibility condition, but only via the term z_g:

Theorem 10. *For given data $f \in L_2(0,T;H^{-1}(\Omega))$, $g \in H^{1/2}(S)$, $u_T \in L_2(\Omega)$ the final value problem (119) is solved by a function $u \in X_1$, whereby*

$$X_1 = L_2(0,T;H^1(\Omega)) \bigcap C([0,T];L_2(\Omega)) \bigcap H^1(0,T;H^{-1}(\Omega)), \tag{154}$$

if and only if the data in terms of (92) and (152) satisfy the compatibility condition

$$u_T - y_f + z_g \in D(e^{-T\Delta_{\gamma_0}}). \tag{155}$$

In the affirmative case, u is uniquely determined in X_1 and has the representation

$$u(t) = e^{t\Delta_{\gamma_0}} e^{-T\Delta_{\gamma_0}}(u_T - y_f + z_g) + \int_0^t e^{(t-s)\Delta} f(s)\,ds - \fint_0^t \Delta e^{(t-s)\Delta_{\gamma_0}} K_0 g(s)\,ds, \tag{156}$$

where the three terms all belong to X_1 as functions of t.

Proof. Given a solution $u \in X_1$, the bijective correspondence yields $u_T = e^{T\Delta_{\gamma_0}} u(0) + y_f - z_g$, so that (155) necessarily holds. Inserting its inversion $u(0) = e^{-T\Delta_{\gamma_0}}(u_T - y_f + z_g)$ into the solution formula from Proposition 9 yields (156); thence uniqueness of u.

If (155) does hold, $u_0 = e^{-T\Delta_{\gamma_0}}(u_T - y_f + z_g)$ is a vector in $L_2(\Omega)$, so the initial value problem with data (f, g, u_0) can be solved by means of Proposition 8. Then one obtains a function $u \in X_1$ that also solves the final value problem (119), since in particular $u(T) = u_T$ is satisfied, cf. the bijection (153) and the definition of u_0.

The final regularity statement follows from the fact that X_1 also is the solution space for the initial value problem in Proposition 8. Indeed, even the improper integral is a solution in X_1 to (128) with data $(f, g, u_0) = (0, g, 0)$, according to Proposition 9; cf. the proof of Lemma 8. Similarly the integral containing f solves an initial value problem with data $(f, 0, 0)$, hence is in X_1. In addition, the first term in (156) is the solution of (128) for data $(0, 0, e^{-T\Delta_{\gamma_0}}(u_T - y_f + z_g))$. ☐

We let Y_1 stand for the set of admissible data. Within $L_2(0,T;H^{-1}(\Omega)) \oplus H^{1/2}(\Gamma) \oplus L_2(\Omega)$ it is the subspace given, via the map $\Phi_1(f, g, u_T) = u_T - y_f + z_g$, as

$$Y_1 = \left\{ (f, g, u_T) \;\middle|\; u_T - y_f + z_g \in D(e^{-T\Delta_{\gamma_0}}) \right\} = D(e^{-T\Delta_{\gamma_0}} \Phi_1). \tag{157}$$

Correspondingly we endow Y_1 with the graph norm of the operator $e^{-T\Delta_{\gamma_0}}\Phi_1$, that is, of the composite map $(f, g, u_T) \mapsto e^{-T\Delta_{\gamma_0}}(u_T - y_f + z_g)$. Again, $e^{-T\Delta_D}\Phi_1(f, g, u_T)$ equals the initial state $u(0)$ steered by f, g to the final state $u(T) = u_T$, as is evident for $t = 0$ in (156).

Recalling that $\mathcal{A} = -\Delta \colon H_0^1(\Omega) \to H^{-1}(\Omega)$, the above-mentioned graph norm is given by

$$\begin{aligned} \|(f, g, u_T)\|_{Y_1}^2 = \|u_T\|_{L_2(\Omega)}^2 + \|g\|_{H^{1/2}(Q)}^2 + \|f\|_{L_2(0,T;H^{-1}(\Omega))}^2 \\ + \int_\Omega \left| e^{-T\Delta_{\gamma_0}} \left(u_T - \int_0^T e^{-(T-s)\mathcal{A}} f(s)\,ds + \fint_0^T \mathcal{A} e^{(T-s)\Delta_{\gamma_0}} K_0 g(s)\,ds \right) \right|^2 dx. \end{aligned} \tag{158}$$

Here the last term is written with explicit integrals to emphasize the complexity of the fully inhomogeneous boundary and final value problem (119).

Completeness of Y_1 follows from continuity of Φ_1, cf. Lemma 8 concerning z_g. Indeed, its composition to the left with the closed operator $e^{-T\Delta_{\gamma_0}}$ in $L_2(\Omega)$ (cf. Proposition 2) is also closed. Hence its domain $D(e^{-T\Delta_{\gamma_0}}\Phi_1) = Y_1$ is complete with respect to the graph norm in (158). As this norm is induced by an inner product when the norm of $H^{-1}(\Omega)$ is taken as $|||\cdot|||_*$ from (26), and when $H^{1/2}(Q)$ is normed as in Remark 17, Y_1 is a Hilbert(-able) space.

Analogously to the proof of Theorem 8, continuity of $(f, g, u_T) \mapsto u$ is now seen at once by inserting the expression $u_0 = e^{-T\Delta_{\gamma_0}}(u_T - y_f + z_g)$ from (153) into the estimate in Proposition 8. Thus we obtain:

Corollary 3. *The unique solution u of problem (119) lying in the Banach space X_1 depends continuously on the data (f, g, u_T) in the Hilbert space Y_1, when these are given the norms in (130) and (158), respectively.*

Taken together, Theorem 10 and Corollary 3 yield that the fully inhomogeneous final value problem (119) for the heat equation is well posed in the spaces X_1 and Y_1.

6. Final Remarks

6.1. Applicability

For the special features of final value problems for Lax–Milgram operators A, it is of course decisive to have a proper subspace $D(e^{TA}) \subsetneq H$, for if $D(e^{TA})$ fills H the compatibility condition (113) will be redundant—and (113) moreover only becomes stronger as the terminal time T increases, if $D(e^{TA})$ decreases with larger T.

Within semigroup theory on a Banach space B, the above means that the ranges $R(e^{t\mathbf{A}})$ should form a *strictly* descending chain of inclusions in the sense that, for $t' > t > 0$,

$$R(e^{t'\mathbf{A}}) \subsetneq R(e^{t\mathbf{A}}) \subsetneq B. \tag{159}$$

Non-strictness is here characterised by the rather special spectral properties of $\mathbf{A}$ in (iv):

Theorem 11. *For a C_0-semigroup $e^{t\mathbf{A}}$ with $\|e^{t\mathbf{A}}\| \le Me^{\omega t}$ the following are equivalent:*

(i) *$e^{t\mathbf{A}}$ is injective and $R(e^{t'\mathbf{A}}) = R(e^{t\mathbf{A}})$ holds for some t, t' with $t' > t \ge 0$.*
(ii) *$e^{t\mathbf{A}}$ is injective with range $R(e^{t\mathbf{A}}) = B$ for every $t \ge 0$.*
(iii) *The semigroup is embedded into a C_0-group $G(t)$ satisfying $\|G(t)\| \le Me^{\omega|t|}$;*
(iv) *The spectrum $\sigma(\mathbf{A})$ is contained in the strip in $\mathbb{C}$ where $-\omega \le \operatorname{Re}\lambda \le \omega$ and*

$$\|(\mathbf{A} - \lambda)^{-n}\| \le M(|\operatorname{Re}\lambda| - \omega)^{-n} \quad \text{for } |\operatorname{Re}\lambda| > \omega,\ n \in \mathbb{N}. \tag{160}$$

Proof. Given (i) for $t > 0$, then $R(e^{(t+\delta)\mathbf{A}}) = R(e^{t\mathbf{A}})$ holds for all $\delta \in [0, t' - t]$ in view of the inclusions (33); and to every $x \in B$ some y satisfies $e^{t\mathbf{A}}e^{\delta\mathbf{A}}y = e^{t\mathbf{A}}x$, which by injectivity gives $x = e^{\delta\mathbf{A}}y$, so that $e^{\delta\mathbf{A}}$ is surjective for such δ. Hence $e^{t\mathbf{A}} = (e^{(t/N)\mathbf{A}})^N$ is a bijection on B with bounded inverse, i.e., $0 \in \rho(e^{t\mathbf{A}})$. If (i) holds for $t = 0$, clearly $0 \in \rho(e^{t'\mathbf{A}})$. In both cases (ii) holds because $0 \in \rho(e^{s\mathbf{A}})$ must necessarily hold for $s > 0$ according to ([23], Theorem 1.6.5), which also states that (iii) holds. (The proof there uses ([23], Lem. 1.6.4) that can be invoked directly from (ii) since the inverse of $e^{t\mathbf{A}}$ is bounded by the Closed Graph Theorem.) Conversely (iii) yields $R(e^{t\mathbf{A}}) = R(G(t)) = B$ and injectivity for all $t \ge 0$, so (ii) and hence (i) holds. That (iv) $\Longrightarrow$ (iii) is part of the content of ([23], Theorem 1.6.3), which also states that (iii) implies (iv) for real λ, but the full statement in (iv) is then obtained from ([23], Remark 1.5.4). $\square$

This result is essentially known, but nonetheless given as a theorem, as it clarifies how widely the present paper applies. Indeed, for V-elliptic Lax–Milgram operators A, the semigroups are uniformly bounded, so $\omega = 0$; thus the strip in (iv) is the imaginary axis $\mathrm{i}\,\mathbb{R}$, but this is contained in $\rho(A)$ by Lemma 4. So except in the pathological case $\sigma(A) = \emptyset$, (iv) will always be violated, as will (i) and (ii). However, since in (i) and (ii) the operator e^{-tA} is injective by Proposition 1, the strict inclusions in (159) hold for $\mathbf{A} = -A$. This proves:

Proposition 10. *For a V-elliptic Lax–Milgram operator A with $\sigma(A) \neq \emptyset$ there is a strictly descending chain of dense domains $D(e^{tA})$ of the inverses $e^{tA} = (e^{-tA})^{-1}$, i.e.,*

$$D(e^{t'A}) \subsetneq D(e^{tA}) \subsetneq H \qquad \text{for } t' > t > 0. \tag{161}$$

Therefore, for elliptic Lax–Milgram operators A with non-empty spectrum, the compatibility condition (113) is without redundancy, and it gets effectively stronger on longer time intervals. Previously, these properties were verified only in a special case in Proposition 7.

Example 1. *It is illuminating to consider the final value problem on $\mathbb{R}^n$, for $\alpha \in \mathbb{C} \setminus \mathbb{R}$,*

$$\partial_t u - \Delta u + \alpha x_1 u = f, \qquad u(T) = u_T. \tag{162}$$

At first glance this might seem to be a minor variation on the heat problem in Section 5, in fact just a zero-order perturbation; and notably a change to $\Omega = \mathbb{R}^n$. However, interestingly it cannot be treated within the present framework: in a paper fundamental to analysis of the Stark effect, Herbst [32] proved for the operator $h(\alpha) = -\Delta + \alpha x_1 I$ with $\operatorname{Im} \alpha \neq 0$ that the minimal realisation $\bar{h}(\alpha)$ also is maximal in $L_2(\mathbb{R}^n)$ with empty spectrum,

$$\sigma(\bar{h}(\alpha)) = \emptyset. \tag{163}$$

Moreover, the numerical range of $h(\alpha)$ itself is an open, slanted halfplane

$$\nu(h(\alpha)) = \{\, z \in \mathbb{C} \mid \operatorname{Re} z > \tfrac{\operatorname{Re} \alpha}{\operatorname{Im} \alpha} \operatorname{Im} z \,\}. \tag{164}$$

Therefore $\bar{h}(\alpha)$ is not sectorial, as $\nu(\bar{h}(\alpha)) \subset \overline{\nu(h(\alpha))}$ shows that (28) does not hold, so existence and uniqueness for the forward problem cannot be derived from Theorem 4. The fact proved in [32] that $e^{-\mathrm{i}t\bar{h}(\alpha)/\alpha}$ is a contraction semigroup, which for $\alpha = \mathrm{i}$ applies to $e^{-t\bar{h}(\mathrm{i})}$ that pertains to (162), entails via the Hille–Yosida theorem the estimate in Theorem 11 (iv) for $-\bar{h}(\mathrm{i})$, but only for $\operatorname{Re} \lambda > 0$. Since $\operatorname{Re} \lambda < 0$ is not covered, it is despite the empty spectrum of $\mathbf{A} = -\bar{h}(\mathrm{i}) = \Delta - \mathrm{i}\, x_1$ not clear whether (159) holds with strict inclusions. Thus it seems open which properties the final value problem (162) for the Herbst operator $h(\alpha)$ can be shown to have.

Remark 18. *Recently Grebenkov, Helffer and Henry [14] studied the complex Airy operator $A = -\Delta + \mathrm{i}\, x_1$ in dimension $n = 1$. They considered realizations defined on $\mathbb{R}_+$ by Dirichlet, Neumann and Robin conditions using the Lax–Milgram lemma, so results on boundary homogenous final value problems for $-\frac{d^2}{dx^2} + \mathrm{i}\, x$ should be straightforward to write down, as in Section 5.1. The study was extended to dimension $n = 2$, under the name of the Bloch–Torrey operator, by Grebenkov and Helffer in [13], where bounded and unbounded domains with C^∞ boundary was treated; in cases with non-empty spectrum there should be easy consequences for the associated final value problems. The realisations induced by a transmission condition at an interface, which was the main theme in [13,14], are defined from a recent extension of the Lax–Milgram lemma due to Almog and Helffer [12], so in this case the properties of the corresponding final value problems are as yet unclear.*

Remark 19. *We expect that extension of the theory to certain systems of parabolic equations with prescribed boundary and final value data should be possible. A useful framework for the discussion of this type of problems could be the pseudo-differential boundary operator calculus, with matrix-formed operators acting in Sobolev spaces of sections of vector bundles, as described in Section 4.1 of [31]. At least the present discussion should carry over to this kind of problems when the realisations called $(P + G)_T$ there are variational, i.e., when they are Lax–Milgram operators for certain triples (H, V, a); this property is analysed in great depth in Section 1.7 of [31], to which we refer the interested reader. It is conceivable that the variational property is unnecessary, and might be avoided using the pseudo-differential boundary operator calculus, but this seems to require an addition to the theory of parabolic systems covered by the calculus in the form of a result on backward uniqueness.*

6.2. Notes

Classical considerations were collected by Liebermann [33] for second order parabolic differential operators (cf. also Evans [34]), with references back to the fundamental L_2-theory including boundary points of Ladyshenskaya, Solonnikov and Uraltseva [35]. A fundamental framework of functional analysis for parabolic Cauchy problems was developed by Lions and Magenes [8]. Later a full regularity theory in scales of anisotropic L_2-Sobolev spaces was worked out for general pseudo-differential parabolic problems by Grubb and Solonnikov [15], who obtained the necessary and sufficient compatibility conditions on the data, including coincidence for half-integer values of the smoothness; cf. also ([31], Theorem 4.1.2). This study was carried over to the corresponding anisotropic L_p-Sobolev spaces by Grubb [36]. A further extension to different integrability properties in time and space was taken up in a systematic study of anisotropic mixed-norm Triebel–Lizorkin spaces on a time cylinder and its flat and curved boundaries by Munch Hansen, the second author and Sickel [27]. Compatibility conditions were addressed for the heat equation in mixed-norm Triebel–Lizorkin spaces in ([27], Section 6.5) and ([29], Chapter 7). In particular, the latter showed that, except for coincidence at half integer smoothness, the recursive formulation of the compatibility conditions in [15] is equivalent to the requirement that the data belong to the null space of a certain matrix-formed operator at the curved corner $\{0\} \times \partial\Omega$. Recent semigroup and Laplace transformation methods were exposed in [30]. Denk and Kaip [37] treated parabolic multi-order systems via the Newton polygon and obtained L_p–L_q regularity results using $\mathcal{R}$-boundedness.

To our knowledge, the literature contains no previous account for pairs of spaces X and Y in which final value problems for parabolic differential equations are well posed.

An early contribution on final value problems for the heat equation was given in 1955 by John [3], who dealt with numerical aspects. In 1961, the idea of reducing the data space to obtain well-posedness was adopted by Miranker [4] for the homogeneous heat equation on $\mathbb{R}$, and he showed that in the space of L_2-functions having compactly supported Fourier transform there is a bijection between the initial and terminal states.

In addition to the injectivity of analytic semigroups in Proposition 1, it is known that $u(0)$ is uniquely determined from $u(T)$ even for t-dependent sesquilinear forms $a(t; v, w)$. This was shown by Lions and Malgrange [16] with an involved argument. It would take us too far to quote the large amount of work on the backward uniqueness in more loosely connected situations, often adopting the log-convexity method (if $|u(t)| \leq |u(T)|^{t/T} |u(0)|^{1-t/T}$ then $u(T) = 0$ implies $u(t) = 0$ for all $t > 0$, hence $u(0) = 0$ by continuity) attributed to Krein, Agmon and Nirenberg. Instead we refer the reader to [38–40] and the references therein.

The method of quasi-reversibility for final value problems was introduced systematically in 1967 by Lattès and Lions [41]. The idea is to perturb the equation $u' + Au = 0$ by adding, e.g., $-\varepsilon^2 A^2$ to obtain a well-posed problem and to derive for its solution u_ε that $u_\varepsilon(x, T)$ approaches u_T for $\varepsilon \to 0$, circumventing analysis of well-posedness of the original final value problem. They assumed $f = 0$ for a V-elliptic self-adjoint A.

Showalter [17] addressed questions that were partly similar to ours. He proposed to perturb instead by $\varepsilon A \partial_t$ under the condition that A is m-accretive with semiangle $\theta \leq \pi/4$ on a Hilbert space for $f = 0$. He claimed uniqueness of solutions, and existence if and only if the final data via the Yosida approximations of $-A$ allow approximation of the initial state. Showalter also identified injectivity of operators in analytic semigroups as an important tool. However, his reduction had certain shortcomings; cf. Remark 1. In comparison we obtain the full well-posedness for general $f \neq 0$ and V-elliptic operators of semiangle $\theta = \operatorname{arccot}(C_3 C_4^{-1})$ belonging to the larger interval $]0, \pi/2[$.

An extensive account of the area around 1975, and of the many previous contributions using a variety of techniques, was provided by Payne [5]. A more recent exposition can be found in Chapters 2 and 3 in Isakov's book [6], and for methods for inverse problems in general the reader may consult Kirsch [42].

In the closely related area of exact and null controllability of parabolic problems, the inequality in Corollary 2 is a little weaker than the observability inequality for the full subdomain $O = \Omega$. In this context, the role of observability was reviewed by Fernandez-Cara and Guerrero [43], emphasising Carleman estimates as a powerful tool in the area. A treatise on Carleman estimates in the parabolic context was given by Koch and Tataru [44].

Author Contributions: Conceptualization, A.-E.C. and J.J.; Method, J.J.; Investigation, A.-E.C. and J.J.; Formal Analysis, A.-E.C.; References, A.-E.C. and J.J.; Original Preparation in Ph.D-Thesis, A.-E.C.; Supervision, J.J.; Editing, J.J. (who also added Section 3.4 and 6).

Acknowledgments: Jon Johnsen was supported by the Danish Research Council, Natural Sciences grant no. 4181-00042. The authors thank H. Amann for his interest and comments on the literature. Also our thanks are due to an anonymous reviewer for indicating the concise proof of Corollary 2.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Christensen, A.-E.; Johnsen, J. On parabolic final value problems and well-posedness. *C. R. Acad. Sci. Paris Ser. I* **2018**, *356*, 301–305.
2. Courant, R.; Hilbert, D. *Methods of Mathematical Physics*; Interscience Publishers, Inc.: New York, NY, USA, 1953.
3. John, F. Numerical solution of the equation of heat conduction for preceding times. *Ann. Mat. Pura Appl.* **1955**, *40*, 129–142.
4. Miranker, W.L. A well posed problem for the backward heat equation. *Proc. Am. Math. Soc.* **1961**, *12*, 243–247.
5. Payne, L.E. *Improperly Posed Problems in Partial Differential Equations*; Society for Industrial and Applied Mathematics: Philadelphia, PA, USA, 1975.
6. Isakov, V. *Inverse Problems for Partial Differential Equations*; Applied Mathematical Sciences; Springer: New York, NY, USA, 1998; Volume 127.
7. Günter, N.M. *Potential Theory and Its Applications to Basic Problems of Mathematical Physics*; Frederick Ungar Publishing Co.: New York, NY, USA, 1967.
8. Lions, J.-L.; Magenes, E. *Non-Homogeneous Boundary Value Problems And Applications*; Springer: New York, NY, USA; Heidelberg, Germany, 1972.
9. Tanabe, H. *Equations of Evolution*; Monographs and Studies in Mathematics; Pitman: Boston, MA, USA, 1979.
10. Temam, R. *Navier—Stokes Equations, Theory and Numerical Analysis*, 3rd ed.; Elsevier Science Publishers B.V.: Amsterdam, The Netherlands, 1984.
11. Amann, H. *Linear and Quasilinear Parabolic Problems*; Monographs in Mathematics; Birkhäuser, Inc.: Boston, MA, USA, 1995; Volume 89.
12. Almog, Y.; Helffer, B. On the Spectrum of Non-Selfadjoint Schrödinger Operators with Compact Resolvent. *Commun. Partial Differ. Equ.* **2015**, *40*, 1441–1466.
13. Grebenkov, D.S.; Helffer, B. On Spectral Properties of the Bloch—Torrey Operator in Two Dimensions. *SIAM J. Math. Anal.* **2018**, *50*, 622–676.
14. Grebenkov, D.S.; Helffer, B.; Henry, R. The complex Airy operator on the line with a semipermeable barrier. *SIAM J. Math. Anal.* **2017**, *49*, 1844–1894.
15. Grubb, G.; Solonnikov, V.A. Solution of parabolic pseudo-differential initial-boundary value problems. *J. Differ. Equ.* **1990**, *87*, 256–304.
16. Lions, J.-L.; Malgrange, B. Sur l'unicité rétrograde dans les problèmes mixtes parabolic. *Math. Scand.* **1960**, *8*, 227–286.
17. Showalter, R.E. The final value problem for evolution equations. *J. Math. Anal. Appl.* **1974**, *47*, 563–572.
18. Janas, J. On unbounded hyponormal operators III. *Stud. Math.* **1994**, *112*, 75–82.
19. Yosida, K. *Functional Analysis*, 6th ed.; Springer: Berlin, Germany; New York, NY, USA, 1980.
20. Grubb, G. *Distributions and Operators*; Graduate Texts in Mathematics; Springer: New York, NY, USA, 2009; Volume 252.
21. Schwartz, L. *Théorie des Distributions*; Hermann: Paris, France, 1966.

22. Hörmander, L. *The Analysis of Linear Partial Differential Operators*; Grundlehren der mathematischen Wissenschaften; Springer: Berlin, Germany, 1983; p. 1985.
23. Pazy, A. *Semigroups of Linear Operators and Applications to Partial Differential Equations*; Applied Mathematical Sciences; Springer: New York, NY, USA, 1983.
24. Reed, M.; Simon, B. *Methods of Modern Mathemtical Physics. I: Functional Analysis*; Academic Press: Cambridge, MA, USA, 1980.
25. Niculescu, C.P.; Persson, L.-E. *Convex Functions and Their Applications. a Contemporary Approach*; CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC; Springer: New York, NY, USA, 2006.
26. Johnsen, J. On spectral properties of Witten-Laplacians, their range projections and Brascamp–Lieb's inequality. *Integr. Equ. Oper. Theory* **2000**, *36*, 288–324.
27. Johnsen, J.; Hansen, S.M.; Sickel, W. Anisotropic Lizorkin–Triebel spaces with mixed norms—Traces on smooth boundaries. *Math. Nachr.* **2015**, *288*, 1327–1359.
28. Johnsen, J.; Sickel, W. On the trace problem for Lizorkin–Triebel spaces with mixed norms. *Math. Nachr.* **2008**, *281*, 1–28.
29. Hansen, S.M. On Parabolic Boundary Problems Treated in Mixed-Norm Lizorkin–Triebel Spaces. Ph.D. Thesis, Aalborg University, Aalborg, Denmark, 2013.
30. Arendt, W.; Batty, C.J.K.; Hieber, M.; Neubrander, F. *Vector-Valued Laplace Transforms and Cauchy Problems*, 2nd ed.; Monographs in Mathematics; Springer: Basel, Switzerland, 2011; Volume 96.
31. Grubb, G. *Functional Calculus of Pseudo-Differential Boundary Problems*, 2nd ed.; Progress in Mathematics; Birkhäuser: Boston, MA, USA, 1996; Volume 65.
32. Herbst, I.W. Dilation analyticity in constant electric field. I. The two body problem. *Commun. Math. Phys.* **1979**, *64*, 279–298.
33. Lieberman, G.M. *Second Order Parabolic Differential Equations*, 2nd ed.; World Scientific Publishing: River Edge, NJ, USA, 2005.
34. Evans, L.C. *Partial Differential Equations*, 2nd ed.; Graduate Studies in Mathematics; American Mathematical Society: Providence, RI, USA, 2010; Volume 19.
35. Ladyzenskaya, O.A.; Solonnikov, V.A.; Ural'ceva, N.N. *Linear and Quasilinear Equations of Parabolic Type*; Translations of Mathematical Monographs; American Mathematical Society: Providence, RI, USA, 1968.
36. Grubb, G. Parameter-elliptic and parabolic pseudodifferential boundary problems in global L_p Sobolev spaces. *Math. Z.* **1995**, *218*, 43–90.
37. Denk, R.; Kaip, M. *General Parabolic Mixed Order Systems In L_p and Applications*; Operator Theory: Advances and Applications; Birkhäuser: Boston, MA, USA, 2013; Volume 239.
38. Dardé, J.; Ervedoza, S. Backward Uniqueness Results for Some Parabolic Equations in an Infinite Rod. Available online: https://hal.archives-ouvertes.fr/hal-01677033 (accessed on 29 March 2018).
39. Hào, D.N.; van Duc, N. Stability results for backward parabolic equations with time-dependent coefficients. *Inverse Probl.* **2011**, *25*, 20. [CrossRef]
40. Kukavica, I. Log-log convexity and backward uniqueness. *Proc. Am. Math. Soc.* **2007**, *135*, 2415–2421.
41. Lattès, R.; Lions, J.-L. *Méthode de quasi-réversibilité et applications*; Travaux et Recherches Mathématiques: Dunod, Paris, 1967.
42. Kirsch, A. *An Introduction to the Mathematical Theory of Inverse Problems*; Applied Mathematical Sciences; Springer-Verlag: New York, NY, USA, 1996; Volume 120.
43. Fernández-Cara, E.; Guerrero, S. Global Carleman inequalities for parabolic systems and applications to controllability. *SIAM J. Control Optim.* **2006**, *45*, 1399–1446.
44. Koch, H.; Tataru, D. Carleman estimates and unique continuation for second order parabolic equations with nonsmooth coefficients. *Commun. Part. Differ. Equ.* **2009**, *34*, 305–366.